Biotechnology Innovation Organization

February 3, 2020

Dockets Management Staff (HFA-305)
Food and Drug Administration
5630 Fishers Lane, Rm. 1061
Rockville, MD 20852

Re: Docket No. FDA-2019-N-4900 Patient-Focused Drug Development Guidance: Incorporating Clinical Outcome Assessments into Endpoints for Regulatory Decision Making; Public Workshop

Dear Sir/Madam:

The Biotechnology Innovation Organization (BIO) thanks the Food and Drug Administration (FDA or Agency) for the opportunity to submit comments regarding FDA's public workshop on Patient-Focused Drug Development Guidance: Incorporating Clinical Outcome Assessments into Endpoints for Regulatory Decision Making.

BIO is the world's largest trade association representing biotechnology companies, academic institutions, state biotechnology centers, and related organizations across the United States and in more than 30 other nations. BIO's members develop medical products and technologies to treat patients afflicted with serious diseases, to delay the onset of these diseases, or to prevent them in the first place.

BIO appreciates FDA's efforts to hold public workshops and develop a series of guidance documents to assist reviewers, drug developers, patient organizations, and other stakeholders with the collection, analysis, and use of patient experience data (PED) for drug development and regulatory decision-making. In general, BIO believes that guidance documents 1-3, and now the discussion guide for Guidance 4, facilitate the advancement and use of systematic approaches to collect meaningful patient and caregiver input that could better inform medical product development and regulatory decision-making. BIO believes that the discussion guide for guidance 4 is clear and concise, includes several instructive examples, and complements and builds upon concepts included in Guidance 3. However, there are several technical topics that would benefit from additional detail and/or case examples in the Discussion Guide for Guidance 4. We have tried to specify what these areas are and where case examples would be helpful. Additionally, across all PFDD guidance and discussion documents there are many connections and interdependencies; BIO believes that the guidance series would greatly benefit from additional cross-referencing between guidance documents.

Because this Discussion Guide will inform the 4th guidance in the series, in addition to providing feedback on Discussion Guide for Guidance 4, we have also provided feedback that may not necessarily need to be addressed in Guidance 4 but that we strongly feel should be addressed in the guidance series so as to best encourage the collection and use of PED for drug development and regulatory decision-making. To this end, BIO has provided in this comment letter suggestions as to additional information and line edits that will allow Guidance 4 to be maximally helpful for incorporating PED in drug development and review in the context of Clinical Outcome Assessments (COAs), as well suggestions as to what information would be helpful for the FDA to address as the guidance document series is finalized.

1201 Maryland Avenue SW
Suite 900
Washington DC 20024

202.962.9200 P
202.488.6306 F
bio.org

BIO Comments on Discussion Guide 4:

I. Guidance 4 should clearly identify the types of regulatory decisions that PED will inform.

As with previous guidance documents in this series, the Discussion Guide for Guidance 4 does not indicate how FDA will be using PED for regulatory decision-making. At the December 6, 2019 public workshop on *Patient-Focused Drug Development: Guidance 4 – Incorporating Clinical Outcome Assessments into Endpoints for Regulatory Decision Making*,[1] Dr. Mullin presented a slide titled "What's Next" which included a figure showing where the PFDD Guidance resides within FDA's Benefit-Risk Framework. However, the Benefit and Risk Management dimension of the framework was blank. While Dr. Mullin spoke to the fact that the table shows COAs inform benefits, and noted that COAs certainly can inform risks, safety, and the burdens associated with a treatment, we recommend that the contents of the table be expanded to reflect the role of COAs in informing both benefits and risks. Additionally, we recommend that this concept is incorporated into this guidance. BIO also recognizes that FDA will be publishing a draft guidance in the first half of 2020[2,3] that will leverage input from the May 16, 2019, public meeting[4] and more comprehensively address the integration of PED into the benefit-risk assessment. Patient priorities regarding avoidance of harm is one area where PED directly feeds into risk and risk management, and we look forward to the forthcoming FDA draft guidance that will address the important topic of how the Agency will incorporate PED data into its risk-benefit decision-making framework. The Discussion Guide indicates PED should be submitted for regulatory decision-making (see lines 18-19, 40-42); however, the guidance is not clear on the types of regulatory decisions for which data will be used.

II. Guidance 4 should address key considerations related to communicating COAs to patients in the label, healthcare providers, payor, and other stakeholders.

FDA has indicated that COAs should be designed with the research question and end-goal in mind. BIO believes that the end-goal includes considerations regarding how the output can be communicated, as it may influence earlier parts of the program design. Both the COA and the endpoint need to be interpretable in the label. BIO requests that FDA expand the guidance to provide the Agency's views on how and where information COAs can be best incorporated in the label and/or be included in other communications to disseminate

---

[1] Public Workshop on Patient-Focused Drug Development: Guidance 4 – Incorporating Clinical Outcome Assessments into Endpoints for Regulatory Decision Making.

[2] 21st Century Cures Act Section 3002(c)(8) "CONTENTS.—The guidance documents described in subsection (a) shall address… how the Secretary, if appropriate, anticipates using relevant patient experience data and related information, including with respect to the structured risk-benefit assessment framework described in section 505(d) of the Federal Food, Drug, and Cosmetic Act (21 U.S.C. 355(d)), to inform regulatory decision making".

[3] *See* FDA Plan for Issuance of Patient-Focused Drug Development Guidance Under 21st Century Cures Act Title III section 3002, May 2017, ("To address Section 3002(c)(8), FDA will issue draft guidance by the end of the second quarter of CY 2020") https://www.fda.gov/media/105979/download.

[4] Public Meeting: Characterizing FDA's Approach to Benefit-Risk Assessment throughout the Medical Product Life Cycle, May 16, 2019, https://healthpolicy.duke.edu/events/benefit-risk-framework-public-workshop.

patient-centric outcomes to other key stakeholders. FDA may also consider publishing assessment reports explaining why certain PROs are/are not included in product labels.

III. Guidance 4 should include reference to major COA development challenges.

BIO suggests that FDA include an additional section in the discussion guide for Guidance 4 that highlights major COA development challenges identified at the concluding panel of the December 6, 2019 meeting.[5] We suggest the Agency provide its current thoughts on each topic by describing the challenges and any best practices, or refer to the appropriate section, to help sponsors navigate these challenges successfully. Below, please find several topics that would be helpful for guidance 4 to address:

A. Developing COAs for heterogeneous populations, including in rare disease populations. This is particularly problematic for trials involving rare diseases where the trial populations are small and there is limited natural history data available;
B. How to develop COAs earlier in development (before phase III) and stressing its importance by identifying critical timepoints and outlining an engagement roadmap;
C. Ways to use COAs in a wider variety of trial types and designs (e.g., open label trials, nonrandomized or nonconcurrent controls (line 805), and the role of modelling (line 715));
D. Personalized endpoints and most-bothersome symptom approach, in light of the recently released draft guidance on migraine and bladder pain[6,7]. Measurement and analytic methods are not quite ready to evaluate these endpoints and may add additional complexity to regulatory and medical decision-making; and
E. How to efficiently address challenges of composite endpoints (e.g., the dual hurdles of achieving a change in the parameter being measured by the endpoint, and then demonstrating that change is clinically meaningful).

IV. Guidance 4 should include additional examples to help illustrate FDA's guidance.

BIO appreciates that FDA provided examples throughout the guidance in order to better illustrate points provided in the Discussion Guide. However, we have identified several additional areas were examples would strengthen the guidance and make FDA's points clearer. BIO would be happy to work with the Agency to collect or develop examples to illustrate the points below:

A. Examples of how to address missing assessments with an explanation of why this approach is preferred. BIO also requests that FDA provide an example of how to address subjects without a baseline measurement and why this is the preferred approach (Lines 112-115).
B. Examples of research questions where removing subjects without baseline measurements may not be a better option (Lines 114-115).
C. Examples of COAs where the domain subsets were adequately developed and validated to measure the subset of domains independently from other domains. It

[5] Public Workshop on Patient-Focused Drug Development: Guidance 4 – Incorporating Clinical Outcome Assessments into Endpoints for Regulatory Decision Making.
[6] Interstitial Cystitis/Bladder Pain Syndrome (IC/BPS): Establishing Effectiveness of Drugs for Treatment Guidance for Industry.
[7] FDA Guidance on Migraine: Developing Drugs for Acute Treatment

would helpful if the guidance included examples beyond item banks like PROMIS or use of briefer instruments like the SF-12 instead of the SF-36 (Line 136).

D. Examples of a multicomponent endpoint and indicate whether it is the same or different than a composite endpoint (Lines 147-149).
E. Examples pertaining to combining different scores from different measurement tools to create a new endpoint using item response theory. It would be helpful if the Agency could provide an example of where a PROMIS tool was used as the basis for a labelling claim (Lines 155-158)
F. Examples of when efficiency (power) may or may not be gained in this situation. An example of gain could be when there is a qualitatively consistent treatment effect across multiple components, whereas an example of loss would be when there are qualitatively inconsistent treatment effects (Lines 169-170).
G. Examples of a multicomponent endpoint FDA considers clinically relevant and interpretable. It would be helpful if FDA could provide examples that demonstrate how these approaches have or could be used successfully (Lines 171).
H. Examples that are more relevant to the clinical trial setting. An example of the 6 Minute Walk Test and "practicing" the activity to qualify for a clinical trial of an experimental treatment in cases of rare disease may be an example that the Agency can use. Another example the Agency could consider is a vision test that is "learned", "practiced", or memorized (Lines 394-398)
I. Examples on how to prevent as well as approach fatigue response if it occurs (Lines 475-479)
J. Examples of a change or disruption (e.g., using paper if electronic device has malfunctioned) and acceptable mitigation from the Agency's perspective (Line 482)

V. Guidance 4 should expand upon several technical sections.

To provide adequate guidance to drug developers on several technical topics included in the discussion guide for Guidance 4, BIO requests that the FDA consider expanding upon these technical sections to include additional information and case examples. While we recognize that guidance documents are not intended to discuss all possible scenarios, the inclusion of more examples would be welcome. The current document does not provide manufacturers with recommendations to some complex issues related to the development of new tools and optimal ways to address problems that may affect the validity of the results. A few specific cases that would benefit from additional examples and additional information, include:

A. Intercurrent events and the "very specific and limited conditions" when "Score generated by the same tool administered ("delivered") via different modes (e.g., interactive voice response; interview; paper-based; electronic device) may be pooled";
B. Developing validated translations of PROs into different languages;
C. Assessing meaningfulness of group level difference for example in the context of the benefit-risk assessment of a novel treatment. Such information may include how between-group differences should be reported, what level of evidence will be considered sufficient, and how FDA will use this evidence to determine treatment benefit and meaningful change; and
D. The section on "Meaningful Within-Patient Change, Other Methods" could be greatly enhanced by the inclusion of additional information on FDA's preference for utilization of methods and context of use. These will not only provide clarity, but

mitigate risk, enhance COA and product development, and ultimately, benefit patients by accelerating bringing products to market.

VI. Guidance 4 should specify what sections of the guidance apply to all phases of drug development and what sections are intended only for late-stage pivotal trials.

Some portions of the guidance appear to have specific applicability for only certain types of trials. For example, in Section C.4. "Defining improvements and worsening", the guidance states on line 302 that "Clinically relevant within-patient thresholds for improvement and worsening should be predefined and justified"; however, early trials are often used to explore the clinically relevant within-patient thresholds and therefore it is not possible to pre-define or justify thresholds at this stage. Therefore, it would be beneficial if the Agency could address the types/phase of trials the guidance is intended to cover, or if applicability varies by section, to call out which sections are applicable to all phases and which are specific to late-stage pivotal trials. In addition, we strongly recommend cross-referencing all the guidance documents in each of the four guidances documents that are part of the PFDD series. This will ensure *all* PFDD principles are applied to COA development and implementation.

VII. Guidance 4 should include an information on implementing the estimand Framework.

BIO requests that the FDA include additional detail in the guidance to address how to implement the estimand framework. The Discussion Guide currently states that plans should be documented in both the protocol and the statistical analysis plan. Traditionally, many of the estimand attributes are included primarily in the statistical analysis plan and not in the study protocols. Details regarding study population, COA analysis, interpretation, and handling of missing data are not necessary for study site personnel to know and they also could lead to misinterpretation of COA instrument data collection. BIO requests that the FDA indicate that information pertaining to the estimand framework be included in the statistical analysis plan.

BIO Comments on the PFDD Guidance Series:

I. The PFDD guidance series should clearly outline how all types of PED will inform regulatory decision-making.

As mentioned above (see comments in Section I. Guidance 4 should clearly identify the types of regulatory decisions that PED will inform), while the guidance documents in the PFDD guidance series have provided important considerations for the collection and use of PED (e.g., Collecting Comprehensive and Representative Input, Methods to Identify What is Important to Patients, Selecting and Developing or Modifying Fit-for-Purpose Clinical Outcome Assessments), to date the guidance documents have not addressed how FDA will be using PED for regulatory decision making. BIO requests that the final PFDD guidance documents address how all types of PED, including patient preference information, will inform regulatory decision-making. BIO requests that FDA address the following points in Guidance 4 as well as the other guidance documents in the series to facilitate efficient

Sponsor design of PED collection and analysis activities in individual development programs:

A. Specify whether the Agency will use PED to inform regulatory decisions such as demonstration of endpoint relevance, benefit-risk evaluation, approvability, labeling, and other decisions via a list of the types of decisions for which FDA will consider PED. For example, note that the guidance currently states that "This document is not directly addressing this integration of benefit and risk, but the methods described can be used to help interpret benefit or risk" (lines 601-603);
B. Indicate and provide any critical considerations or evidentiary standards that may be unique to a particular decision type (see examples decision types above);
C. Map the decision type to the relevant guidance in the series where that topic is discussed;
D. Indicate areas where FDA expects to issue guidance in the future (e.g., FDA's plan for addressing the use of PED in in the integration/evaluation of benefit-risk[8]); and
E. Provide direction for how sponsors can most efficiently engage the Agency to discuss PED topics, especially those not yet covered in guidance (e.g., benefit-risk evaluation, labeling, see also Section II. of this comment letter).

Additionally, to support the importance of PED data to inform regulatory decision-making, it would be helpful for the FDA to provide more detailed and consistent information on how each type of PED was either submitted with the application or by another stakeholder, and how it was considered for regulatory decision-making through the Statement of Patient Experience Checklist and in the product approval review documents such as the Multi-Disciplinary Review or Clinical Review.

II. The PFDD guidance series should identify opportunities for early interaction between FDA & Sponsors.

In the PFDD guidance document series, FDA indicates that "FDA encourages stakeholders to have early interactions with FDA and to obtain feedback from the relevant FDA review division when considering collection of PED related to the burden of disease and burden of treatment". However, the Agency does not outline which meeting pathways stakeholders should use to discuss such data with the Agency. Additionally, there are currently challenges with securing meetings with FDA.[9] For successful incorporation of patient input into drug development, BIO requests that FDA outline appropriate meeting pathways for companies to obtain early feedback and outline the types of information that should be presented to FDA prior to and in such meetings. As relevant to Guidance 4, it would be helpful for FDA to indicate what meeting types should be used for the discussion of patient-driven endpoints. BIO also requests that FDA consider this comment in the context of the entire guidance series, including reference to meeting pathways for stakeholders to discuss PED at the different stages of the drug development lifecycle.

---

[8] *See* FDA Plan for Issuance of Patient-Focused Drug Development Guidance Under 21st Century Cures Act Title III section 3002, May 2017, ("To address Section 3002(c)(8), FDA will issue draft guidance by the end of the second quarter of CY 2020") https://www.fda.gov/media/105979/download

[9] PDUFA Reauthorization Performance Goals and Procedures Fiscal Years 2018 through 2022.

BIO appreciates this opportunity to submit comments regarding FDA's Patient-Focused Drug Development Guidance: Incorporating Clinical Outcome Assessments into Endpoints for Regulatory Decision Making; Public Workshop. BIO would be happy to assist the Agency in developing case examples in order to help strengthen and make more useful Guidance 4 of the PFDD guidance series. We would also be pleased to provide further input or clarification of our comments, as needed.

Sincerely,

/S/
Danielle Friend, Ph.D.
Senior Director, Science and Regulatory Affairs
Biotechnology Innovation Organization

SPECIFIC COMMENTS

SECTION	ISSUE	PROPOSED CHANGE
I. INTRODUCTION		
A. Guidance Series		
B. Document Summary		
Lines	In the discussion guide there is no reference to health literacy; however, health literacy and numeracy should be a foundation of all patient measures to achieve the following: a. Greatest inclusion of patient populations, including diverse populations (who are statistically more likely to have low health literacy) b. Reduced likelihood of patient non-participation in clinical trial, due to lack of understanding c. If health literate, numerate measures are used in the study, health literate language will then be a part of patient labeling for new molecules or indication, which will facilitate broad patient understanding *(supports line 760 that this will be foundational to labeling claims)* d. It is important to include people from a range of health literacy levels in the testing of the proposed patient measures *(including*	BIO requests that FDA reference the importance of health literacy and include reference to other resources, including: • Center for Disease Control and Prevention Health Literacy: Development and Test Materials[10] • Institute of Medicine of the National Academies: Health Literacy Principles: Guidance for Making Information Understandable, Useful, and Navigable[11] • National Academy of Medicine: Strategies to Enhance Numeracy Skills[12]

[10] Centers for Disease Control and Prevention Health Literacy: Development and Test Materials.
[11] Institute of Medicine of the National Academies: Health Literacy Principles: Guidance for Making Information Understandable, Useful, and Navigable
[12] National Academy of Medicine: Strategies to Enhance Numeracy Skills

SECTION	ISSUE	PROPOSED CHANGE
	alternate measures line 438), to make sure they are widely understood.	
II. ESTIMAND FRAMEWORK OVERVIEW		
Line 58, entire section on estimand framework		BIO recommends that the Agency note that obtaining patient stakeholder input on the estimand model and/or COA framework will help ensure the programs is designed to measure what matters most to patients.
Line 58	In this section the FDA indicates that "*An estimand is a quantity used to define a treatment effect in a clinical study;*" however, it is unclear what is meant by "quantity." Additionally, the provided estimand components relate to the draft ICH E9(R1) addendum and do not align with the final version, for example, reference to "treatment" is missing.	BIO requests that the FDA provide clarity as to what is meant by "quantity." BIO also requests that FDA ensure that this section aligns with ICH work.
Line 59	In this section the FDA indicates that "*Protocols should specify intercurrent events and how they will be accounted for in analyses to address the scientific question of interest.*" However, protocols should specify intercurrent events and how they will be accounted for in the analysis to address the scientific questions of interest.	BIO requests the following edit: "*Protocols* ***and/or statistical analysis plans*** *should specify intercurrent events and how they will be accounted for in analyses to address the scientific question of interest.*"
A. COA Research Objective: Foundation for Your Work		
Lines 61-62	In this section the FDA indicates that "*The attributes listed above should be clearly defined prior to developing a protocol and included in both*	BIO requests that the FDA indicate that intercurrent events should be included in the protocol and/or Statistical Analysis Plan.

SECTION	ISSUE	PROPOSED CHANGE
	the protocol and Statistical Analysis Plan (SAP). They will determine the data collected, procedures, and other sections of the protocol beyond statistical methods. The attributes also drive the SAP and communication of trial results, as highlighted in Figure 1."	We recommend the following revision to ensure the same level of detail that is expected with other endpoints is included in the protocol and that there is a balance of what goes in the protocol versus the statistical analysis plan. For example, specific details about the types of analyses planned should be included in the statistical analysis plan: "The attributes listed above should be clearly defined prior to developing a protocol and included in both the protocol and Statistical Analysis Plan (SAP). **Relevant information should be included in the protocol. However, specific details about the types of analyses planned should be included in the SAP. Information on intercurrent events should be included in the protocol and/or SAP.**"
Line 66 and figure 1	Figure 1: Attributes of an estimand placed in context • Intercurrent Events • Population-Level Summary	For clarity, BIO requests that FDA replace "estimand" with "estimand framework" in the figure 1 title and inside the figure.
Line 70	In this section, FDA indicates that "*The essence of clinical research is to ask important questions and answer them with appropriate studies (ICH E8(R1)).*"	We believe that there may be a typo in this section. To this end, BIO requests the following edit: The FDA indicates that "*The essence of clinical research is to ask important questions and answer them with appropriate studies (ICH E**9**~~8~~(R1)).*"
B. Target Study Population: In Whom Are You Going to Do the Research and Which Subject Records Are in the Analysis?		
Lines 82-115, entire section	This section is meant to discuss choice of patient population for the estimand. However, it discusses analysis sets, intent-to-treat, and missing data, which are all elements of trial conduct and/or	BIO requests that FDA consider creating a separate section which discusses considerations for selecting the target population.

SECTION	ISSUE	PROPOSED CHANGE
	analysis, none of which are required or relevant for defining an estimand.	
Lines 94-95, Table 1	This table outlines considerations for defining a COA target study population.	BIO requests that FDA clarify whether these definitions refer to "analysis sets" rather than to true "populations" in the sense of the ICH E9 addendum.
Lines 97-100	In this section, FDA indicates that "*For sponsors considering an effectiveness claim from a COA-derived endpoint in a randomized trial, the intent-to-treat (ITT) population generally should be used to preserve the benefits of randomization.*" One of the anticipated benefits of the ICH E9 addendum is that it provides a framework that allows stakeholders to discuss various strategies to handle an intercurrent event.	BIO requests that FDA formulate the draft guidance using ICH E9 addendum language (i.e., explicitly using the strategy(ies) introduced in the E9 addendum). We also request that FDA clarify whether ITT should be considered a feature of the *estimand* or rather the *estimator*. We request that FDA clarify if there is a strong preference for the "treatment policy", as it is implied in this sentence. Finally, we request that FDA provide an example of when it would be acceptable to use a target population other than intent-to-treat (ITT) for an effectiveness claim based on a COA-derived endpoint. BIO also request the following edit: "*For sponsors considering an* ~~***effectiveness***~~ ***efficacy*** *claim from a COA-derived endpoint in a randomized trial, the intent-to-treat (ITT) population generally should be used to preserve the benefits of randomization.*" Additionally, while the merits of recommending full follow-up of data are clear and useful to include in this document, this paragraph on missing data is ultimately about data collection and analysis (i.e., estimation), rather than about the patient population that the trial is attempting to estimate something about. BIO suggests that reference to data collection and analysis be included in a separate section.

SECTION	ISSUE	PROPOSED CHANGE
Lines 103-105	In this section FDA indicates that "*Consider how interpretation of the COA-derived endpoint changes if all patients in a trial are not eligible for the COA.*"	For clarity, BIO requests the following edits: "*Consider how interpretation of the COA-derived endpoint changes if* ***not*** *all patients in a trial are* ~~***not***~~ *eligible for the COA.*"
Line 110	In this section the FDA indicates that "*Every effort should be made to have high completion rates throughout the study.*"	BIO requests what is considered "high completion" rates throughout the study.
Lines 112-115	In this section, FDA indicates that "*Because there is the potential for patients to have missing assessments, sponsors should clearly specify in the SAP how missing observations will be dealt with for clear interpretation. Removing subjects without a baseline measurement is common but depending on the research question it may not be the better option.*"	BIO requests the following edit: "Removing subjects without a baseline measurement is common but depending on the research question it may not be the better option. **In the example provided in Appendix 2, if all data were missing for the baseline assessments, the data collected during screening replaced the missing baseline assessments. This was particularly critical given the small sample size in the Phase 3 study for this rare condition.**"
C. Endpoint of Interest: What Are You Testing or Measuring in the Target Study Population?		
1. Endpoint definition(s)		
126-128	In this section the FDA indicates that "*Hence, assessment of an endpoint's reliability, content validity, construct validity, as well as ability to detect change are important (refer to FDA PFDD G3 Public Workshop Discussion Document for details).*"	BIO requests the following edit: "*Hence, assessment of* ~~***an endpoint's***~~ ***a COA's*** *reliability, content validity, construct validity, as well as ability to detect change are important (refer to FDA PFDD G3 Public Workshop Discussion Document for details).*"
Line 127	In this section FDA refers to FDA PFDD G3 Public Workshop Discussion Document for details.	BIO requests that when the draft guidance is developed FDA provide the complete reference with year and hyperlink to

SECTION	ISSUE	PROPOSED CHANGE
		resources (e.g., transcript, etc.) for "FDA PFDD G3 Public Workshop."
Line 131	Digital tools are not discussed or defined in the text. Given the movement to capture patient data using digital tools in clinical trials, or to facilitate decentralized trials to reduce burden on patients, BIO suggests that FDA include definitions for these terms.	BIO suggests adding the following definition after line 131, or in a footnote, as appropriate. **Digital Endpoint: A precisely defined variable intended to reflect an outcome of interest that is statistically analyzed to address a particular research question derived from data captured with digital health technology tool.** BIO also suggests adding a definition of digital COA here, to formally acknowledge another type of clinical outcome assessment: a digital monitoring clinical outcome assessment. FDA has already used the term "digital monitoring clinical outcome assessment" in their COA qualification submission decisions for various tools such as ActiMyo®5 and Physical Activity Accelerometry Assessment for Analgesic Clinical Trials (PAACT)6. Consistent with the previously used term, suggest including the following definition: Digital Monitoring Clinical Outcome Assessment (dmCOA) - A type of clinical outcome assessment. A measurement based on a report that comes from technology after the detection and measurement of activity/function, behaviors, or other manifestations related to a disease or condition. dmCOA measures typically do not require the patient to actively perform a standardized task as in the case of a performance outcome assessment. Rather, they can be obtained passively as the patient goes about their daily life and activities in non-clinical settings (e.g., passive monitoring of falls or sleep

SECTION	ISSUE	PROPOSED CHANGE
		quality using wearable instruments). Please see also comments on line 489.
Line 136	In this section FDA indicates "*The use of domain subsets to support clinical trial endpoints assumes the COA was adequately developed and validated to measure the subset of domains independently from the other domains.*"	BIO requests that the FDA provide additional clarity regarding what is meant by "adequately developed and validated to measure the subset of domains."
Line 139	In this section the FDA indicates that "*A complex, multi-domain claim cannot be substantiated by instruments that do not adequately measure the individual components of the domain.*" Please add the following: One example is the Asthma Daytime Symptom Diary and the Asthma Nighttime Symptom Diary. The FDA only qualified 3 core categories of asthma symptoms for use. Other relevant and important concepts are also assessed but are not considered part of the diary.	BIO requests that the FDA make the following addition: "***One example is the Asthma Daytime Symptom Diary and the Asthma Nighttime Symptom Diary. The FDA only qualified 3 core categories of asthma symptoms for use. Other relevant and important concepts are also assessed but are not considered part of the diary.***"
2. Pooling different tools and/or different concepts to construct the endpoint?		
Lines 151-152	Other analytical methods, such as global tests, could potentially be used to pool scores from different tools of a similar type, e.g., ***patient-reported outcomes*** (PROs).	BIO requests that FDA clarify what is meant by "pooled scores from different tools."
Line 175	In this section FDA indicates "*For some rare diseases with heterogeneous patient populations and variable disease manifestations, it may be challenging to assess a single concept of interest across all patients.*"	BIO believes that the Discussion Guide would be significantly strengthened if the discussion of how to develop endpoints for heterogeneous patient populations for rare diseases were expanded. The discussion at the FDA public meeting on December 6, 2019 also highlighted other sources of heterogeneity that should be included in the Guidance document. The guidance would benefit from discussion

SECTION	ISSUE	PROPOSED CHANGE
		regarding how to have enough people with a particular manifestation to show a statistically significant improvement versus capturing rare manifestations that may be important to patients.
Lines 183-184	In this section FDA indicates that "*An example dichotomous (event) endpoint is the multidomain responder index (MDRI) approach, which thus far has not been demonstrated as a viable approach based on evidence submitted to FDA.*"	BIO requests that the FDA provide a reference for "multidomain responder index (MDRI) approach. BIO also requests that FDA provide suggestions on how to address heterogeneity in rare diseases and possibly a COA that represents a viable approach to addressing heterogeneity instead of providing examples of what is not adequate.
Lines 203-204	In this section FDA indicates that "*Similar concerns exist with personalized or individualized endpoints, which often are analyzed descriptively as exploratory endpoints.*"	BIO requests that FDA clarify whether "personalized or individualized endpoints" refer to situations where each subject may not get the same question.
Lines 202-211		While FDA describes concerns with personalized endpoints in this section, BIO believes that it would benefit from expansion and clarification in general. Specifically, we recommend that FDA expand and clarify whether the section is addressing measuring a patient's most bothersome symptom. Additionally, we also recommend that the Agency define and clarify "personalized endpoint" and include it in the PFDD glossary of terms as it may differ among patients and may change throughout the trial. It may also be helpful to reference the report from the Duke Margolis public workshop on personalized endpoints, which includes discussion of some of the considerations for implementing these types of endpoints.[13]

[13] Duke Margolis Meeting on Developing Personalized Clinical Outcome Assessments

SECTION	ISSUE	PROPOSED CHANGE
		Examples of personalized endpoints, such as included in the FDA's guidance for developing drugs for acute treatments[14] would also be very helpful to include in the draft guidance. Furthermore, BIO agrees with the Agency's position that the process to construct a personalized endpoint should be standardized, and that the same set of outcome assessments should be assessed for all patients, regardless of their own personalized endpoint to allow for an assessment of any new or worsening symptoms and/or functional limitations during the trial duration. However, individualized endpoints confer multiple benefits in describing clinically important change—and it would be helpful to understand if there are ways to limit FDA concerns when outcomes are put in appropriate context. We therefore recommend that the Agency address and clarify the following: 1. First, individualized endpoints are inherently 'patient-centric', so with adequate construction, changes are arguably inherently clinically important. 2. Second, in certain degenerative disorders like AD, floor and ceiling effects are problematic over time, and cognitive performance measures lack ecological validity. Individualized endpoints can help address some of the challenges of interpreting change. 3. Third, if the individualized endpoint requires change over the course of the trial, it is an additional opportunity to describe 'worsening' or 'improvement' that, again, are arguably of more importance than other outcomes assessed and may otherwise have not been

[14] FDA Guidance on Migraine: Developing Drugs for Acute Treatment

SECTION	ISSUE	PROPOSED CHANGE
		captured. There is a role if administered appropriately and agreed to a priori as in an adaptive design.
Lines 236-242	In this section FDA indicates that "*Although scores yielded by different modes are generally considered to be comparable when there is no difference between modes in terms of the wording of item stems and response options, item formats, the appearance and usage of graphics or other visuals, or order of the items (see FDA PFDD G3 Public Workshop Discussion Document for further discussion), administering a tool using more than one mode or method per study can introduce noise (i.e., construct-irrelevant variance in COA score) that may not be completely random and may make it more difficult to discern treatment effects.*"	For clarification, BIO requests the following edit: "*Although scores yielded by different modes are generally considered to be comparable when there is no difference between modes in terms of the wording of item stems and response options, item formats, the appearance and usage of graphics or other visuals, or order of the items (see FDA PFDD G3 Public Workshop Discussion Document for further discussion), administering a tool using more than one mode or method per study can introduce* ~~***noise***~~ ***variability*** *(i.e., construct-irrelevant variance in COA score) that may not be completely random and may make it more difficult to discern treatment effects.*"
3. Timing of assessments		
Line 289	In this section FDA indicates that "*The timing of anchor scale administration should align with both the recall period and the administration of the corresponding COA (e.g., patient global impression of severity (PGIS) with PRO timing; clinician global impression of severity with ClinRO timing). "In this section FDA uses the term "anchor scale."*	BIO requests that the Agency include in future guidance a definition for the term "anchor scale" as it is not currently defined in the discussion guide. This is a new requirement and may present difficulties, particularly for legacy PRO instruments and anchors such as patient global impression scales. Many patient global impression scales ask respondents to respond based on how they feel "right now" or "today." To align an anchor such as a PGI-S to the COA instrument's recall period may not be appropriate or feasible.
4. Defining improvements and worsening		
5. Clinical trial duration and COA-based endpoints		

SECTION	ISSUE	PROPOSED CHANGE
D. Intercurrent Events: What Can Affect Your Measurement's Interpretation?		
1. Use of assistive devices, concomitant medications, and other therapies		
Lines 324-325	In this section FDA indicates that "*While missing data is a part of the definition [of intercurrent events]", it is not the only definition.*"	Per ICH E9(R1) missing data is not part of the definition, indicating that "*Because the estimand is to be defined in advance of trial design, neither study withdrawal nor other reasons for missing data (e.g., administrative censoring in trials with survival outcomes) are in themselves intercurrent events.*" BIO requests that the Discussion Guide be revised accordingly. BIO also requests that reference to the handling of missing data in this section (intercurrent events) be moved to a separate section.
Lines 340-341	In this section FDA indicates that "*Case report forms (CRFs) for data collection should include information on whether an assistive device (and what type) was used during the test.*" Some intercurrent events can be defined as a part of protocol development and analysis planning (e.g., subsequent therapy, rescue medication, use of an assistive device); however, others (e.g., patient in a trial breaks their leg in a car accident) will be difficult to collect and plan for in the CRF a priori.	In general, the CRF should map the data necessary for collection in order to appropriately capture the compliance to the treatment strategy (not only assistive devices) defined in the corresponding estimand attribute. We request FDA to clarify and provide more detail on the overall expected intercurrent events Sponsors should collect, and the content of the CRF.
Line 343-354	In this section the FDA provides two examples of intercurrent events.	BIO requests that the Discussion Guide provide detail as to what should be done when currently validated COAs do not explicitly account for intercurrent events like the examples given. It would be helpful for the Agency to elaborate as to whether the COAs need to be modified and revalidated in such instances. BIO also requests that FDA provide examples of how to account for intercurrent events in the design and analysis of the data.

SECTION	ISSUE	PROPOSED CHANGE
2. Impact of disease/condition progression, treatment, and potential intercurrent events		
Lines 362-366	This section does not reference health literacy.	Having health literate, numerate items built in from the beginning will reduce unnecessary cognitive load and improve the duration and accuracy of patient responses. BIO requests that the FDA reference health literacy in this section.
Lines 367-376	In this section FDA indicates that "*Missed or incomplete assessments due to disease progression or treatment side effects "may provide meaningful information on the effect of a treatment and hence may be incorporated into a variable [(or endpoint)], with appropriate summary measure, that describes a meaningful treatment effect" (ICH E9(R1))*." Since model-based estimates generally tend to be "very sensitive" to model misspecification, it is recommended that supplementary and sensitivity analyses be conducted to examine how much the results/findings change under various assumptions about the missing data mechanism (National Research Council, 2010)." It is useful to create a variable for missing or incomplete assessments due to disease progression. It is also useful to conduct sensitivity analyses. Currently, it is not clear as to whether the reader is to assume the statistical analysis plan should explicitly specify how the sensitivity analyses are to be interpreted.	BIO recommends that this be included in the guidance document as it is not simply a matter of noting sources of competing risks as indicated in Line 390.
Lines 389-391	In this section FDA indicates that "*Because changes in cognitive and physical function may still*	BIO requests that the FDA clarify what "note" means in this section. In earlier sections the document suggests that all

SECTION	ISSUE	PROPOSED CHANGE
	occur during the study, it is important to note sources of competing risks and other intercurrent events in the SAP and Study Report."	intercurrent events should be listed in the protocol AND SAP. This sentence says SAP and Study Report. BIO requests that, in this section and throughout the document, the FDA note where this information should be included.
3. Practices effects		
Line 421-449	In this section FDA indicates "*Some general strategies for mitigating practice effects are summarized below. These strategies may be used in isolation but may be more effective when used in combination."*	BIO finds the suggested general strategies for mitigating practice effects to be useful recommendations and would encourage FDA to retain in the future guidance document.
Line 435	In this section FDA indicates that "*Having a long run-in period allows large practice effects to occur for the first few assessments until its magnitude does not significantly increase such that the baseline and postbaseline score are minimally affected by practice effects.*" However, increasing the length of the run-in period and implementing multiple assessments during the run-in period may not be feasible in the context of clinical trial. For example, increased patient burden due to multiple site visits or prolonged run-in period may lead to progression of the disease (e.g., cancer) while withholding potential treatment.	BIO requests that FDA provide a reference supporting the adjacent approach and an example of implementing this approach in the context of a clinical trial.
Lines 438-444	This section does not reference health literacy and numeracy.	BIO requests that FDA indicate that if alternate forms are used, make sure they are also health literate and numerate.
4. Participant burden		
Line 451-460	This section discusses issues pertaining to participant burden. BIO believes that participant	BIO requests that FDA consider changing the language to soften how COAs are perceived to be "burdensome" especially

SECTION	ISSUE	PROPOSED CHANGE
	burden should be discussed in the context of total trial procedures and assessments (e.g., the number of COAs and total number of questions administered as well as the frequency of administration). Additionally, there is a negative connotation associated with the word "burden," which is further highlighted in the manner in which this concept is addressed in the Discussion Guide. This section also does not reference health literacy and numeracy.	since there is no empirical evidence to suggest a threshold for burden. Additionally, BIO requests that this section acknowledge the totality of trial participation burden and suggest mitigation options acceptable to the Agency. We also request that the Agency expand this section to include a discussion of use of Clinical Trial Sponsor-supplied handheld devices to patients and/or caregivers, including for regulatory decision-making. Additionally, we request that this discussion address FDA's positioning regarding validity of the data collected from above-mentioned sources as the differences in device operating systems could possibly change the visual appearance/layout of COA questions and potentially influence the responses (including any practice effects). BIO requests that FDA indicate that health literate, numerate forms can help to reduce participant burden, particularly for people with lower levels of health literacy.
5. Mode of administration		
Lines 489-492	In PFDD Guidance 2, the agency mentions two terms, "method of administration" and "mode of administration."	BIO requests that FDA clarify whether "mode" refers to the medium used to capture data (e.g., paper, electronic device), and "method" refers to the COA respondent and/or administrator of the COA. BIO also suggests that FDA note that electronic approaches can also include data gathered from passive monitoring devices.
6. Missing data and event driven COA reporting		
Lines 499-506	In this section, FDA indicates that "*Programming errors can result in significant amounts of missing data which impedes interpretation of analysis results. For example, a COA may be designed to give patients the option to report additional events and event-related symptoms not reported during*	This section describes a specific programming error that would result in an event-triggered eCRF page not appropriately opening and therefore causing missing data. The example is very specific to software failure; however, there are many other kinds of programming errors that would result in similar missing data situations.

SECTION	ISSUE	PROPOSED CHANGE
	the day; however, a potential programming error could cause the additional questions to not be administered at the end of the day. Large amounts of missing data would be generated, resulting in underreporting of the event and the study endpoint itself being unreliable and uninterpretable."	We also request that the Agency provide additional detail regarding how Sponsors may address missing data, beyond the example within the context of software errors. For example, this section should discuss the nature of missing data and how it introduces bias and analytical methods, such as whether missing data occurs across the treatment groups randomly, and factors that are associated with missing value (e.g., age, gender, and disease severity), and guidance how to treat each situation differently. Reference to data monitoring and data management plans and reports as well as site training and patient education materials should be noted and discussed as means to address missing data and improve overall data quality.
7. Missing scale-level data		
Lines 507-521	In this section FDA discusses the case when the patient omits all items of a domain. Often time responses exist except for an item or two. Missing data (e.g., patient forgot to complete PRO) should be distinguished from data that do not exist or data that are not considered meaningful due to an intercurrent event (e.g. patient discontinued treatment). Missing data should be addressed in SAP but not in protocol, as this could easily be misinterpreted by study site personnel.	BIO also requests a discussion as to how to handle those cases (for many scoring manuals, single missing items are allowed, and imputed based on the non-missing items) which may be exemplified by examples.
E. Population-Level Summary: What is the Final Way All Data Are Summarized and Analyzed?		
1. Landmark analysis		
Lines 535-539	In this section FDA indicates that "*Sponsors should justify the use of and time in which a landmark*	BIO requests that "landmark analysis" be clearly defined within the document and included in the Glossary.

SECTION	ISSUE	PROPOSED CHANGE
	analysis (an analysis at a fixed time point, e.g. 12 weeks) is to be performed. If a COA-based endpoint is collected repeatedly, information may be lost in conducting a landmark analysis. However, even when conducting a landmark analysis at a fixed time point, data from intermediate time points (i.e., measurements taken prior to the fixed time point) can still be included in the model."	Additionally, in many cases a landmark analysis may not be the best approach and may be underpowered. To this end, BIO requests that the Agency include details in the guidance discussing alternatives, including approaches such as Area Under the Curve, slope analysis, Time x treatment interaction, among others.
Lines 539-542	In this section FDA indicates that "*Interpretation of an analysis of overall COA score over time may be difficult in the presence of missing data. The interpretation of potential analyses when COA data collection is truncated due to death or other events should be carefully discussed within the research team.*" This section refers to missing data in the context of a death of a patient. However, "when COA data collection is truncated due to death" the lack of COA data is not generally considered missing data. According to the National Academy of Science's 2010 report on The Prevention and Treatment of Missing Data in Clinical Trials[15] and its distillation[16], "Missing data are defined as values that are not available and that would be meaningful for analysis if they were observed."	BIO recommends that FDA consider expanding this paragraph to state: "**Measures of quality of life or COAs are usually not meaningful for patients who have died and hence would not be considered as missing data under this definition[5,6].**" BIO requests that FDA ensure that the language in the Discussion Guide aligns with the language of ICH E9(R1) on addressing intercurrent events and request that sponsors clearly define intercurrent events in the protocol and/or statistical analysis plan. We recommend this be cross-referenced to section D.7.

[15] National Academies of Science, Engineer, and Math. The Prevention and Treatment of Missing Data in Clinical Trials. 2010.
[16] Little et al., The Prevention and Treatment of Missing Data. (2012) New England Journal of Medicine. 367:1355-1360.

SECTION	ISSUE	PROPOSED CHANGE
	Thus, measures of quality of life or COAs are usually not meaningful for patients who have died and hence would not be considered as missing data under this definition.	
Lines 544-548	This section provides details around the analysis of ordinal data, but ordinal scales are very difficult for people with low health literacy.	BIO requests that FDA consider language rather than numbers for each option. For instance, instead of a 1-5 scale, consider repeating the following language as options for each question: Strongly disagree, disagree, neutral, agree, strongly agree.[17]
Line 556	In this section FDA indicates that "for example, estimates may differ if death is considered a deterioration event versus censored."	BIO requests that the Agency provide details as to under what circumstances it would be best to include death as a deterioration event versus censored.
Lines 562-581	This section focused on responder analysis and percent change from baseline. However, in this section, a distinction should be made between an ordinal or a continuous COA-based endpoint used as such in a main analysis and, if justified, converting this endpoint into a binary outcome (e.g., responder, no responder) as an adjunct, descriptive analysis solely for the purpose of enriching clinical interpretation and meaning (on the other hand). Responder analysis is best positioned as a descriptive display and as an augmentation to – as a complement and supplement to – the main	While BIO finds this statement to be useful and requests that it be retained in the future guidance, we suggest FDA avoid a complete rejection of responder type analyses and percent change analyses. The general rejection of responder type analyses and percent change analyses seems too undifferentiated and partially contradictory to other parts of the document and FDA endpoint recommendations for some disease areas. Additionally, the position outlines does not appear to be consistent with the FDA Guidance on Patient-Reported Outcome Measures: Use in Medial Product Development to Support Labeling Claims[18] and Discussion Guide for Guidance 2 and 3, Patient-Focused Drug Development Guidance: Methods to Identify What is Important to Patients and Select, Develop or Modify Fit-for purpose

[17] Holbrook, Allyson & Cho, Young & Johnson, Timothy. (2006). The Impact of Question and Respondent Characteristics on Comprehension and Mapping Difficulties. Public Opinion Quarterly - PUBLIC OPIN QUART. 70. 565-595. 10.1093/poq/nfl027

[18] FDA 2009 Guidance on Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims

SECTION	ISSUE	PROPOSED CHANGE
	analysis based on the full original scale of measurement using established statistical methods (e.g., repeated measures or random coefficient models when the data are longitudinal). Such a responder analysis is not intended to replace the main analysis based on the original metric of COA-based endpoints.	Clinical Outcome Assessments (page 18, "From a regulatory standpoint, FDA is more interested in what constitutes a meaningful within-patient change in scores from the patient perspective …")[19] BIO requests that whenever possible, consistency between this guidance and other guidance documents should be emphasized. Responder analyses are often included as supportive evidence to try to answer whether the difference between treatment groups is large enough to demonstrate whether sufficiently more subjects benefited from treatment. Additionally, other groups such as IQWIG have recommended specific meaningful change thresholds such as >=20% change from baseline.
2. Analyzing ordinal data		
3. Time-to event analysis		
4. Responder analyses and percent change from baseline		
Lines 581-584	In this section FDA indicates that "*Strange occurrences arise, for example in randomized withdrawal studies we have seen subjects needing to reach a percent change from baseline threshold who end up needing significantly higher symptom burden to go back on treatment compared to symptom levels needed to enter the trial based on inclusion criteria.*"	BIO requests that FDA include this sentence as a broader example to be added to the appendix instead of this paragraph. We also request that FDA clarify what is meant by "strange occurrences."
III. MEANINGFUL WITHIN-PATIENT CHANGE		
Line 595	In this section FDA indicates that "*To aid in the interpretation of study results, FDA is interested in what constitutes a meaningful within-patient*	BIO requests that FDA expand this section to include information as to whether both anchor- and distribution-based methods should be used in support of a triangulation approach.

[19] Discussion Guide for Guidance 2 and 3, Patient-Focused Drug Development Guidance: Methods to Identify What is Important to Patients and Select, Develop or Modify Fit-for purpose Clinical Outcome Assessments.

SECTION	ISSUE	PROPOSED CHANGE
	change (i.e., improvement and deterioration from the patients' perspective) in the concepts assessed by COAs. Statistical significance can be achieved for small differences between comparator groups, but this finding does not indicate whether individual patients have experienced meaningful clinical benefit."	
Line 596	In this section FDA indicates "*Anchors selected for the trial should be plainly understood in context, easier to interpret than the clinical outcome itself, and sufficiently associated with the target COA and/or endpoint."*	BIO requests that the Agency provide guidance on the type of evidence the Agency would like to justify what would be determined to be "sufficiently associated" with the target COA and/or endpoint. Some researchers have advocated a correlation coefficient of at least 0.30.[20] BIO also requests that FDA indicate that for some progressive diseases "no change" or a "reduced rate of deterioration" may also be considered meaningful. Additional detail on how to use individual patient changes to determine meaningful clinical benefit would also be helpful. Finally, BIO suggests removing "Technical Summary" from the title as it does not include all elements listed in this section.
Lines 604-607	In this section FDA indicates that "*To aid in the interpretation of the COA-based endpoint results, sponsors should propose an appropriate threshold(s) (e.g., a range of score change) that would constitute a clinically meaningful within-patient change in scores in the target patient*	BIO requests that FDA provide details such as: • The Agencies proposed timing for review; • Whether a non-a priori definition acceptable (i.e. estimating these thresholds) with pivotal data prior to unblinding; and • Whether data should be pooled across arms.

[20] Revicki, D., Hays, R.D., Cella, D., Sloan, J. (2008). Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. JOunral of Clinical Epidemiology. Feb.; 61(2):102-109.

SECTION	ISSUE	PROPOSED CHANGE
	population for FDA review;" however it is unclear as to whether the Agency has a proposed timing for this review.	
Lines 604-607	In this section FDA indicates that "*To aid in the interpretation of the COA-based endpoint results, sponsors should propose an appropriate threshold(s) (e.g., a range of score change) that would constitute a clinically meaningful within-patient change in scores in the target patient population for FDA review.*"	BIO requests that FDA provide additional detail as to how multiple thresholds be used operationally to interpret trial results.
Lines 609-614	In this section FDA indicates "*In addition, if the selected threshold(s) are based on transformed scores (e.g., linear 610 transformation of a 0-4 raw score scale to a 0-100 score scale), it is important to consider score interpretability of the meaningful change threshold(s) for both transformed scores and raw scores;*" however there is additional information that would help guide Sponsors and that was not included in this section.	BIO requests that FDA clarify and specify the different types of proposed score transformations that are suitable and useful for the evaluation and interpretation of clinically meaningful change. Such information may include: • What specific score transformations would result in selected threshold(s) based on transformed scores being less than a one-category change on the raw score scale; • What specific score transformations would result in transformed scores being at least a one-category change on the raw score scale; and • Where does the direct linear function transformation stand [where there is a direct one-to-one mapping between transformed and untransformed (raw) score].
Lines 616-628	In this section FDA indicates that "*FDA is more interested in what constitutes a meaningful within-patient change in scores from the patient perspective (i.e., individual patient level).*" As it is currently written, it is unclear under what circumstances FDA would prefer within-patient	Although "between-group" change is not appropriate for developing interpretation thresholds for individual patients, it remains an important metric to inform whether the change between treatment groups is clinically meaningful. Firstly, between-group effects smaller than the within-patient threshold can represent large differences in the number of patients meeting the within-patient criteria due to between patient

SECTION	ISSUE	PROPOSED CHANGE
	change in scores versus the approach taken in the section above on responder analysis and percent change from baseline.	variability. Secondly, even if the between-group benefits were evenly applied across patients, effects that are less meaningful on an individual level can produce meaningful benefit at a population level. A discussion clarifying the difference between within-patient interpretation thresholds (i.e., meaningful change estimates for change in at the individual level) and between-group interpretation thresholds (i.e., meaningful change estimates at the group-level) would be beneficial. Clarifying that both are of import but for different reasons (i.e. estimating interpretation thresholds for individual change vs. estimates of meaningful difference in the point estimation). Additional information including how between-group differences should be reported, what level of evidence will be considered sufficient for use of between-group differences, and how FDA will use this evidence to determine treatment benefit and meaningful change would greatly strengthen the Discussion Guide. Further, it would be helpful for the Discussion Guide to describe how meaningful within-patient change and between group differences are used separately and in combination for regulatory decision making.
Lines 625-627	In this section FDA indicate that "*The terms minimally clinically important difference (MCID) and minimum important difference (MID) do not define meaningful within-patient change if derived from group-level data and therefore should be avoided.*"	There is a lot of confusion within the clinical community regarding the differences between MCID, MID, and MCT. Often these terms are used inter-changeably. BIO requests that FDA more explicitly define these concepts and consider including them in an updated version of the glossary. It would also be helpful for the Agency to clarify recommendations on how to determine MCID/MID versus within-patient changes. For example what methodology does the Agency recommend for determining if between-group differences are meaningful.
A. Anchor-Based Methods to Establish Meaningful Within Patient Change		

SECTION	ISSUE	PROPOSED CHANGE
Lines 637-642	In this section FDA indicates that "*The anchor measure(s) are used as external criteria to define patients who have or have not experienced a meaningful change in their condition, with the change in COA score evaluated in these sets of patients. Sponsors should provide evidence for what constitutes a meaningful change on the anchor scale by specifying and justifying the anchor response category that represents a clinically meaningful change to patients on the anchor scale, e.g., a 2-category decrease on a 5-category patient global impression of severity scale.*"	As it reads, this section implies that sponsors should provide a rationale to support what subjects consider a meaningful category change on the anchor. BIO requests that FDA clarify which methodology should be used to determine what amount of category change in an anchor is meaningful. While Section B and C focus on supplemental methods, section A lacks a clear recommendation and description of the primary anchor method. Additional clarity regarding the type of evidence needed to justify the anchor category representing meaningful change would be helpful.
Lines 646-647	In this section FDA indicates that "*Selected anchors would be plainly understood in context, easier to interpret than the COA itself, and sufficiently associated with the target COA or COA endpoint.*"	BIO requests that FDA clarify what is meant by "sufficiently associated with the target COA or COA endpoint." BIO requests FDA indicate in the Discussion Guide that they will accept other outcomes such as caregiver burden as an anchor.
Line 658	In this section FDA indicates that "*Well-established clinical outcomes (if relevant)" is an anchor that is sometimes recommended to generate appropriate threshold(s) that represent a meaningful within-patient change in the target patient population."*	BIO requests that FDA include examples of possible relevant clinical outcomes in the Discussion Guide.
B. Using Empirical Cumulative Distribution Function and Probability Density		

SECTION	ISSUE	PROPOSED CHANGE
Lines 663-692 and entire section	The described eCDF and PDF curves in this section are only descriptive displays of the data and do not provide a clear approach as to how to define meaningful within-patient change.	BIO supports FDA's inclusion of this information from the 2009 guidance. To further increase utility, we recommend the following: 1. Providing example eCDFs and PDFs with interpretations; 2. Clarifying why PDFs provide more intuitive overviews of the shape, dispersion, and skewness of the distribution of the change from baseline in the endpoints; and 3. Discussions as to how diagnostics ultimately inform clinically meaningful change. A holistic interpretative context to the example figures will consolidate this section further. BIO believes that it would be valuable to add a discussion on what considerations should be taken when determining a meaningful within patient change (MWPC) cutoff based on anchor-based data.
Lines 676-677	In this section FDA indicates that "*This should be considered when choosing an anchor summary and interpreting these figures and data.*"	BIO requests that FDA expand this section to include information on how these should be considered when choosing an anchor.
Lines 689-692	In this section FDA provides an example of density function curves of change in COA score from baseline to primary time point by change in PGIS Score.	BIO requests that additional detail be added to this example especially as to how to translate the graphs into a clinically meaningful difference and providing this additional level of detail in future guidance would make the example more instructive.
C. Other Methods		
Lines 699-700	In this section, FDA indicates that "*The qualitative research methods in the PFDD Guidance 1 and Guidance 2 documents are frequently used, including cognitive interviews, exit interviews, or surveys to help inform the improvement threshold.*"	BIO recommends that the Agency clarify that cognitive interviews in the context of ClinROs refer to clinicians (e.g., physicians, nurses).

SECTION	ISSUE	PROPOSED CHANGE
Lines 708-712	In this section FDA indicates that "*Distribution-based methods (e.g., effect sizes, certain proportions of the standard deviation and/or standard error of measurement) do not directly take into account the patient voice and as such cannot be the primary evidence for within-patient clinical meaningfulness. Distribution based methods can provide information about measurement variability.*"	While it is clear that anchor-based estimates are favored, FDA also states that this is not always possible to do (e.g., lack of suitable anchor). BIO requests the FDA provide additional details on which distribution-based approaches and other approaches are favored when anchor-based estimates are not possible, or as a complement to anchor-based approaches in a triangulation approach to the development of interpretation thresholds,[21] as well as the perceived strengths and limitations of these approaches.
Lines 715-717, page 22	In this section FDA indicates that "*Unless there is significant knowledge about how a COA performs in a specific context of use, FDA does not recommend using receiver operator characteristic (ROC) curve analysis as a primary method to determine the thresholds for within-patient meaningful change score.*"	BIO appreciates the concerns with using ROC curves as the primary method for determining the MWPC threshold; however, we see the value of its use in determining the appropriateness of the anchor in much the same way as eCDFs. We suggest that FDA acknowledge the ROC curves as a valid alternative presentation to eCDF visualization. BIO also requests that FDA clarify which two groups the agency is referring to (e.g., two treatment groups). In general, ROC should be conducted on all treatment groups combined so that the threshold derived is treatment agnostic.
Lines 717-722, page 22	In this section FDA indicates that "*The ROC curve method is a model-based approach, such that different models may yield different threshold values. Additionally, the ROC curve method is partially a distributional-based approach, such that the distribution of the change scores of the two*	We suggest FDA amend the text to better highlight the utility of ROC curves in the *process* of identifying a suitable threshold rather than focusing on its use in *defining* the threshold.

[21] Leidy, NK., Wyrwhich KW. (200). Bridging the Gap: Using Triangulation Methodology to Estimate Minimal Clinically Important Differences (MCIDs). Journal of Chronic Obstructive Pulmonary Disease. March 2(1): 157-165.

SECTION	ISSUE	PROPOSED CHANGE
	groups will determine the location of the threshold. The most sensitive threshold identified by ROC may not actually be the most clinically meaningful threshold to patients."	
1. Potentially useful emerging methods		
2. Distribution-based methods		
3. Receiver operator characteristic curve analysis		
Line 727	In this section FDA provides information regarding within-patient change to clinical trial data. The discussion guide seems incomplete as it does not link the established meaningful difference to power planning. For instance, it is not clear as to whether FDA will require that non-meaningful differences be excluded from the confidence interval, to enable approval / labelling claims.	BIO requests that as FDA is developing the guidance in addition to enabling interpretation of observed data, the established meaningful difference should also be used as the delta in power planning.
Lines 742-743 and Figure 4	In this section the Discussion Guide indicates that "*The treatment effect occurs in the range patients consider to be clinically meaningful.*"	BIO requests that the FDA clarify what is meant by "range" is referenced here.
Lines 754-755		BIO requests that the FDA clarify what is considered "not in range" and how the range is defined.
IV. ADDITIONAL CONSIDERATIONS		
Lines 760-794		If an endpoint is not a key secondary or in the endpoint hierarchy, how much of the work suggested in this document is necessary or recommended? PFDD implies that COA endpoints can support regulatory decision making even if they aren't in the endpoint hierarchy.

SECTION	ISSUE	PROPOSED CHANGE
Lines 762-789		BIO requests that FDA clarify whether the numerical subheadings (1 to 16) are meant to be in order of priority.
Lines 765-766	In this section FDA indicates that "*COAs intended to support meaningful outcomes to patients (i.e., labeling claims or other communications) are fit-for-purpose and sensitive to detect clinically meaningful changes*" should be confirmed when planning a study.	BIO requests FDA to clarify whether "clinically meaningful changes" refer to "within-patient clinically meaningful changes." BIO also requests that FDA provide additional detail regarding what "other communications" refers to.
Line 770	In this section the FDA indicates that "*How blinding or masking will be implemented (e.g., assessor blinding)*" should be considered when planning a study.	BIO requests that FDA clarify whether the Agency considers "blinding" and "masking" as interchangeable.
Line 773	In this section FDA indicates that "*Procedures for training are well-described*" should be considered when planning a study.	As this is the first mention of training, additional information on who is being trained, what is meant by "well-described," and where procedures should be described would be helpful.
Line 774	In this section FDA indicates that "*Content and scoring information are clearly delineated in the clinical trial protocol*" should be considered when planning a study.	Scoring information is relevant for the statistical analysis plan and may be misinterpreted and confusing to study site personnel if include in the study protocol.
Line 776	In this section FDA indicates that "*COA-based endpoints intended to support approval and/or labeling claims are appropriately positioned in the endpoint testing hierarchy*" should be considered when planning a study. COA endpoints intended to support approval or to inform benefit risk do not necessarily need to be positioned in the testing hierarchy (e.g., PRO-CTCAE and tolerability). COA data that provide information about the patient	To support the importance of these COA data as well, it may be helpful for the FDA to ensure comment on these data. It would be helpful to have additional guidance on the difference between supporting approval and labeling claims in this context.

SECTION	ISSUE	PROPOSED CHANGE
	experience may not be positioned in the endpoint testing hierarchy, if not intended for labeling.	
Lines 786-787	In this section FDA indicates that when planning a study, one should consider "*description of how between-group differences will be portrayed (e.g., cumulative distribution function)*"	BIO requests that the Agency consider clarifying in what context the Agency is interested in between-group differences.
Line 775	In this section FDA indicates that "*Plans for COA scoring are consistent with those used during instrument development*" should be considered when planning a study.	BIO suggests replacing "development" with "psychometric validation" or "confirmed during the instruments psychometric validation".
Line 778	In this section FDA indicates that "*Plans for multiplicity adjustment*" should be considered when planning a study.	BIO requests that FDA consider the following edit: "*Plans for multiplicity adjustment* ***if the COA is to support labeling.*** **Please refer to FDA guidance on "Multiple Endpoints in Clinical Trials Guidance**" BIO also suggests inclusion of additional detail regarding adjusting for multiple comparison. BIO requests that FDA cross-reference to the FDA guidance on Multiple Endpoints in Clinical Trials.[22]
Lines 788-789	In this section FDA indicates that "*Data collection, data storage, and data handling and transmission of procedures, including electronic COAs, are specified*" should be considered when planning a study.	BIO requests that FDA clarify what is meant by "transmission of procedures". BIO also requests that FDA provide additional detail as to where are these items should be specified.

[22] FDA Guidance on Multiple Endpoints in Clinical Trials.

SECTION	ISSUE	PROPOSED CHANGE
Lines 791-793	In this section FDA indicates that "*Both SPIRIT (Calvert et al, 2018) and CONSORT (Calvert et al, 2013) consensus documents have been published with extensive details on what PRO information should be included in trial protocols and manuscripts*."	BIO requests that FDA replace "SPIRIT" with "SPIRIT-PRO" and "CONSORT" with "CONSORT-PRO" for clarity.
A. Other Study Design Considerations		
B. Formatting and Submission Considerations		
APPENDIX 1: CASE STUDY OF ESTIMAND FRAMEWORK		
Entire section	The Appendix 1 case study seems simplistic and not representative of oncology clinical trials. For example, missing data is not addressed in the example, but is often an issue in the oncology setting.	BIO requests that the example in Appendix 1 include missing data as well as information on how the issue of missing data can be addressed.
A. Example Research Objective		
1. Define COA scientific research questions a priori		
2. Define target study population based on the research question a priori		
3. Define endpoint of interest based on the research question a priori		
Lines 923-924	In this section FDA indicates that "*We are looking at Week 28, which is around a 6-month time point in which the cumulative effects of the product in terms of both efficacy and toxicity have equilibrated.*"	For the endpoint definition, it is not clear why a single time point is relevant from the patient perspective. This is a concern, in particular, when there is potentially a higher toxicity burden at the start of treatment for an experimental drug given as an add-on therapy to standard of care. BIO requests FDA clarify the relevance of a single time point from the patient perspective.
4. Address intercurrent events in alignment with the research question		

SECTION	ISSUE	PROPOSED CHANGE
Line 943, entire section		We ask the agency to reference strategies, described in the language of the ICH E9 addendum, with which the intercurrent events are handled.
Line 948	In this section FDA indicates that "*For patients who discontinue treatment, progress, start physical therapy, initiate subsequent therapy or experience any other intercurrent event, we continue to collect physical function assessments regardless of these intercurrent events and will include them in our analysis.*"	BIO requests that FDA provide an example with the rationale for the preferred approach as to how manufacturers should account for intercurrent events in the analysis of the data.
Line 953, Column: "handling of intercurrent event", line 959	This section indicates that "*Table 3 presents a list of additional intercurrent events that may impact interpretation of physical function.*"	BIO requests the Agency provide additional guidance in this diagram on how to handle the intercurrent events. Table 3 could be more informative if specific examples of ways to handle intercurrent events were provided. Currently this section states that handling of intercurrent events should be prespecified; however, there are multiple ways to handle intercurrent events but the guidance does not provide this detail. An example that could be added is as follows: the impact of hospitalization on mobility - different ways to handle this intercurrent event and how the different ways of handling this type of intercurrent event affect the interpretation of the results such as eCOA and scoring of mobility.
5. Define population-level summary based on research question a priori		
6. Prespecify statistical analysis plan		
Line 988	The section indicates that the Statistical Analysis Plan should include a description of why the intercurrent events are included in the statistical analyses.	BIO requests that FDA provide additional details as to whether there should be a description of a sensitivity analysis should be conducted and how the results should be interpreted.

SECTION	ISSUE	PROPOSED CHANGE
Lines 996-997		BIO requests that FDA make clearer in the guidance how to best handle the death of a patient. Even if a low proportion of deaths are expected, once death is defined as an intercurrent event, then a strategy on how to handle it needs to be defined. BIO recommends the Agency clarify the strategy to handle such cases.
Line 1001	In this section FDA indicates that "*Suitable supplementary analyses should be performed to challenge the assumptions of the prespecified analysis by incorporating reasons for missingness in the analysis.*"	It is not clear as to what is meant by "suitable" and "incorporating reasons for missingness in the analysis." Sponsors want to conduct "suitable" sensitivity analyses to evaluate the robustness of the results. Often, it may not be known for certain why data is missing. Sponsors can estimate that there may be data missing because people are in the placebo group and withdraw from the study to investigate other treatments or data could be missing for an illness unrelated to the treatment given in the study.
B. Summary of Decisions made in This Case Study		
Line 1010, Table 5: Summary of Estimand Decision Made		We understand that all intercurrent events but death are to be handled using a "treatment policy" strategy. We ask the agency to clearly state early in the document the way to handle any kind of intercurrent events and provide enough details for a clear differentiation.
APPENDIX 2: EXAMPLE FROM GENE THERPAY		
Line 1017, entire section		It would be helpful to tie this example back to the estimand framework and discuss how the PerfO data was used in the regulatory decision-making process. A detailed description of evidence needs, limitations, and interpretation would significantly strengthen the Discussion Guide.
Lines 1073	In this section FDA indicates that "*After a phase 1 trial, the sponsor identified the need to develop a*	BIO requests that FDA provide additional details as to why FDA accepted the MLMT tool for a labelled indication.

SECTION	ISSUE	PROPOSED CHANGE
	novel clinically meaningful PerfO endpoint specific to the treatment effect of Luxturna on the target patient population, and went on to develop and validate the MLMT, in discussion with FDA. This illustrates the importance of carefully designed and conducted early-phase trials in informing the design of late-phase trials."	
Lines 1106		BIO requests that FDA provide additional details regarding how the clinically meaningful benefit in functional vision was determined using anchored based methods would be helpful in future guidance. BIO also requests that FDA add another bullet to explain why an MLMT score change of ≥2 is considered a clinically meaningful benefit in functional vision between Lines 1112 and 1113 as follows: "**The sponsor initially proposed an MLMT score change of ≥ 1 as clinically meaningful. However, the FDA believed that an MLMT score change of 1 may represent a background fluctuation in both the treatment and control groups based on the frequency distribution of MLMT score changes in the control group. They advised the sponsor to regard a score change of 2 as clinically meaningful.**"
APPENDIX 3: references		
APPENDIX 4: GLOSSARY		
Line 1240, Glossary	Several definitions included in the glossary are not aligned with the descriptions in the main document (e.g., estimand).	BIO requests that the FDA ensure that definitions included in the glossary align with those included in the main body of the document.

SECTION	ISSUE	PROPOSED CHANGE
Lines 1240-1245 on		BIO requests that the FDA develop a health literate glossary of these terms so that sponsors describe these terms the same way in future patient labeling (reducing cognitive burden and unnecessary confusion for patients).