February 28, 2022

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

Re: Docket No. FDA–2021-D-1146: Real-World Data: Assessing Registries to Support Regulatory Decision-Making for Drug and Biological Products

Dear Sir/Madam:

The Biotechnology Innovation Organization (BIO) thanks the Food and Drug Administration (FDA or Agency) for the opportunity to submit comments regarding the draft guidance on **Real-World Data: Assessing Registries to Support Regulatory Decision-Making for Drug and Biological Products.**

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 recommends the following overarching comments:

Collaborating with registry holders/partners

The draft guidance refers to registries that may be initiated by the drug sponsor, as well as registries initiated by a third party, such as an academic institution or other registry holder. For the latter, some of the suggested activities and information may not be applicable or available to the drug sponsor. For example, a drug sponsor working with an outside registry holder may be able to audit the registry holder's processes for data quality but not usually the data themselves due to ethical and/or data privacy reasons. On the other hand, if the drug sponsor is also the registry holder, such audits on source data could be performed by the drug sponsor. Therefore, BIO recommends that the guidance further clarify the expected "sponsor" activities throughout, whether the activities are intended for the registry holder only or the drug sponsor, and that consideration be given to those scenarios where a drug sponsor may not be able to access or audit certain information due to contractual/legal reasons. BIO recommends that the Agency consider holding a workshop or public meeting with relevant stakeholders to further discuss the aforementioned items.

Scope of Guidance (Types of Registries)

The draft guidance seems to address traditional registries where patients are consented and there is a data collection/validation plan. However, it also mentions registries used for quality purposes at Health Care Organizations (HCOs, line 82). BIO believes this is referencing Qualified Clinical Data Registries (QCDRs) that have been created to address CMS quality requirements and consist of HCO-submitted EMR data (e.g., RISE, Axon). While labeled as

registries, there is no consent, enrollment, data collection plan, etc., associated with these QCDRs. There are also data providers who have developed "registries" from integrating multiple sources of secondary data such as claims/EMR/PRO (again, without consent, enrollment, data collection plan, etc.). BIO recommends that the Agency clarify whether or not such integrated data 'registries' are in scope for this guidance.

BIO recommends that this draft guidance document should address registries where patients are consented and enrolled for a specific purpose with a data collection plan (Line 139) while integrated data "registries" mentioned above should be out of scope (they are more appropriately covered within FDA's EHR/claims guidance). The term "registry" is being used broadly in industry by sponsors and data providers, and this document could help to clarify what registry data would be considered fit-for-purpose by FDA in supporting regulatory decision-making.

BIO recommends that the Agency clarify the use of data from foreign registries and if there are specific considerations for Special Population registries. BIO also recommends that the Agency consider providing additional detail about how sponsors could quantify or at least qualify some of the key concepts mentioned in the guidance, such as accuracy and completeness. Lastly, it may be helpful to reference the AHRQ manual on creating registries for outcomes if FDA recommends this manual as a resource: https://effectivehealthcare.ahrq.gov/products/registries-guide-4th-edition/users-guide .

It would also be helpful if the guidance further discussed on and described scenarios when registry data could be used to support pre-marketing applications versus when registry data could be used to support post-approval applications. In addition, it would be helpful if the guidance were modified to include recommended elements for a safety/Adverse Event registry (e.g., grade, diagnosis).

Harmonization of RWE Guidance

BIO recommends that the Agency consider providing more consistency on the topics covered across guidance documents. For example, consider including additional considerations for hypothesis testing and validation in the registries guidance as these topics are covered extensively in the EHR/claims guidance. Alternatively, FDA may explain in more depth why considerations may be more important for one data source versus another. Topics that are likely relevant across RWD sources such as claims, EHRs and registries (e.g., validation, linkage, hypothesis testing, provenance, etc) are not treated equally across the guidance documents, which may be a source of confusion for sponsors. Alignment across guidance documents (to the extent possible) would be welcomed by sponsors.

Similarly, it would be helpful for FDA to clarify and reference the relationship between this draft guidance and the *Assessing EHR and Medical Claims Data* draft guidance. BIO recommends that the Agency consider clarifying the need for following the *Assessing EHR and Medical Claims* guidance when linking secondary data to the registry (Line 385). Adding a reference to the EHR/claims draft guidance document seems appropriate in sections referencing linkage with these types of data. Similarly, when applicable, BIO recommends that the Agency reference the draft guidance on *FDA Data Standards for Drug and Biological Product Submissions Containing Real-World Data*. BIO also encourages the Agency to collaborate with the European Medicines Agency (EMA), as they have released guidelines for registry studies for regulatory decision-making within Europe. And, similar to the EMA Guideline, it would be helpful if FDA provided

more definitions upfront in the guidance, e.g., provide a clear definition for a registry versus registry based clinical study.

Sincerely,

/s/
Camelia Thompson, Ph.D.
Senior Director, Science and Regulatory Affairs
Biotechnology Innovation Organization

SPECIFIC COMMENTS

SECTION	ISSUE	PROPOSED CHANGE
I. INTRODUCTION		
Lines 20-24	Lines 20-24 seem to suggest that only studies intended to support new indications (e.g., sNDA) for approved drugs or studies to satisfy postapproval study requirements are in scope. However, registry data on the current standard of care could play a role in studies supporting a new drug application, for example, used to create an external control arm for a pivotal single-arm trial. Excluding such studies from the scope of this guidance may be a missed opportunity.	BIO recommends the Agency clarify whether the scope of the guidance includes the use of registries to support new drug applications (NDAs) and Biologics License Applications (BLAs).
Lines 26-28	The draft guidance states, "This guidance provides sponsors and other stakeholders with considerations when either proposing to design a registry or using an existing registry to support regulatory decision-making about a drug's effectiveness or safety." The regulatory use of RWD is primarily foreseen to support new indications of already approved drugs (as outlined in lines 21-24), however, there might be cases when data from registries can also support initial NDA/BLA. It is suggested not to limit the scope to new indication applications or post-approval requirements.	BIO recommends the following edits: "This guidance provides sponsors and other stakeholders with considerations when either proposing to design a registry or using an existing registry to support regulatory decision-making about a drug's effectiveness or safety **for initial or supplemental applications.**"
Lines 40-41	The draft guidance states, "RWE is the clinical evidence about the usage and the potential benefit or risks of a medical product derived from analysis of RWD." This definition might suggest that RWE only pertains to a product and doesn't capture natural history designs or evidence related to disease states/indications that is commonly provided by registries.	BIO recommends the following edit to define RWE as pertaining to a medical product or disease state across all of the RWD guidance documents: "RWE is the clinical evidence about the usage and the potential benefit or risks of a medical product **or about a disease state** derived from analysis of RWD."

SECTION	ISSUE	PROPOSED CHANGE
Lines 37 to 41	The draft guidance could elaborate on the definitions of RWD and RWE with definition included on the website https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence	BIO recommends the Agency provide links to this website as it provides good examples of RWD and RWE that can help the readers understand the difference.
Lines 54-57	The draft guidance states, "Whether registry data are fit-for-use in regulatory decision-making depends on the attributes that support the collection of relevant and reliable data (described in this guidance) as well as additional scientific considerations related to study design and study conduct that are beyond the scope of this guidance." It would be helpful for the Agency to provide reference to the additional scientific considerations.	BIO recommends that the Agency provide references for the cited additional scientific considerations, for example, to other FDA guidance, whether already available or planned.
II. BACKGROUND		
Entire Section	The background section should contain additional examples of registries that were used for regulatory decision making.	BIO recommends adding examples of registries used to generate evidence for regulatory decision-making. For example, the Nordic ITP registry was used to describe early ITP outcomes in patients under standard of care (McGrath et al. 2021, https://pubmed.ncbi.nlm.nih.gov/34416023/). Additionally, the EXPECT pregnancy registry examined the occurrence of major congenital anomalies in pregnant women with severe asthma treated with omalizumab (Namazy et al. 2019, https://www.jacionline.org/article/S0091-6749(19)30690-6/fulltext). BIO notes that there are many examples where registries were used to support decision making and perhaps the best examples for educational purposes would be examples where sponsors could access the summary basis of approval documents.

SECTION	ISSUE	PROPOSED CHANGE
Lines 70-73	Registry owners might find it useful to adapt code lists or questionnaires over time. However, this could impact interpretability of the data.	BIO recommends that the Agency mention that in longitudinal registries it is strongly recommended to apply consistent data standards across timepoints.
Lines 70-79	The draft guidance states, "For the purposes of this guidance…" This reads as if only curated registries are in scope of this guidance.	BIO recommends the Agency clarify if the scope of the guidance is the re-use of existing sponsor registries for a different purpose than those already evaluated and/oruse of registries from other sources than the sponsor.
Lines 72-73	The term "enrolling patient" in these lines seems to point to prospective data collection. However, the mention of external data sources such as medical claims, EHRs, curation, and linkage in subsequent lines (Lines 74-79) seem to point to retrospective collection.	BIO recommends that the Agency clarify whether the draft guidance includes both prospective data collection and retrospective data collection. The guidance could benefit from separate sections to cover information for both scenarios.
Lines 76-78	The draft guidance states, "Such external data sources can include data from medical claims, from pharmacy and/or laboratory databases, and from EHRs, blood banks, and/or medical device outputs." Vital statistics databases/indexes are missing from this list.	BIO recommends the following edit: "Such external data sources can include**, but not be limited to,** data from medical claims, from pharmacy and/or laboratory databases, and from EHRs, blood banks, **vital statistics databases/indexes,** and/or medical device outputs."
Line 76	The draft guidance does not include other types of links like genomic data.	BIO recommends that the Agency include other types of links like genomic data.
Lines 87-88	Consider referencing relevant FDA guidance on the design and curation of registries.	BIO recommends that the Agency consider referencing relevant FDA guidance on the design and curation of registries.
Lines 92-115	<u>Potential Uses of Registries</u> The potential uses listed in this section seem to reflect a perspective of registries generally being used to support interventional studies. We wish to highlight the	BIO recommends the Agency consider including the following potential uses of registries of characterization of treatment patterns/sequencing, examining the financial and humanistic burden of disease, assessment of clinical outcomes in

SECTION	ISSUE	PROPOSED CHANGE
	importance of registries in generating insights outside the context of interventional study support, such as in treatment sequencing, financial/humanistic burden of disease, or even comparative effectiveness studies. There are also differences in uses for disease registries and treatment registries, whereby the former tends to focus on patient characterization, treatment patterns, and long-term clinical outcomes, and the latter tends to focus on safety.	identifying underserved populations, and comparative effectiveness when an interventional study is infeasible or untimely. BIO also recommends providing examples where registries are used to collect patient data in the context of a randomized trial.
Lines 95 - 115	While it is acknowledged that the section provides only examples where registries have the potential to support medicine development, pragmatic randomized trials may represent an important example of interventional studies using registries to support inferences about effectiveness or safety.	BIO recommends that the Agency clarify whether pragmatic randomized controlled trials using a registry to collect some (or all) patient outcome data fall within the scope of this guidance.
Lines 95 - 115	Some of the listed purposes seem to be mainly a sponsor's risk (e.g., sample site planning for future studies), whereas others have a regulatory dimension (especially the last one: use of external control arm).	BIO recommends that the Agency clarify whether FDA feedback is recommended for all of these purposes.
Lines 109-110	The draft guidance states, "Supporting, in appropriate clinical circumstances, inferences about safety and effectiveness in the context of: …"	BIO recommends the Agency add a **third** contextual sub-bullet under the bullet point in lines 109 - 110: "Supporting, in appropriate clinical circumstances, inferences about safety and effectiveness in the context of:" - **Bridging clinical outcomes to an underrepresented sub-population or alternative standard-of-care** BIO also requests additional examples of what the Agency considers "appropriate clinical circumstances".

SECTION	ISSUE	PROPOSED CHANGE
Line 115	This line mentions registries may support "An externally controlled trial including registry data as an external control arm" and, further on, a definition for externally controlled trial is given. While it is acknowledged that the section provides only examples where registries have the potential to support medical product development, randomized controlled trials with a hybrid control arm (Gray et.al. 2020; Schmidli et.al. 2020) can be regarded as having better quality of evidence compared to externally controlled trials. Gray CM, Grimson F, Layton D, Pocock S, Kim J. A framework for methodological choice and evidence assessment for studies using external comparators from real-world data. Drug safety. 2020 Jul;43:623-33. Schmidli H, Häring DA, Thomas M, Cassidy A, Weber S, Bretz F. Beyond randomized clinical trials: Use of external controls. Clinical Pharmacology & Therapeutics. 2020 Apr;107(4):806-16. Mishra-Kalyani PS, Amiri Kordestani L, Rivera DR, Singh H, Ibrahim A, DeClaro RA, Shen Y, Tang S, Sridhara R, Kluetz PG, Concato J, Pazdur R, Beaver JA. External Control Arms in Oncology: Current Use and Future Directions. Ann Oncol. 2022 Jan 10:S0923-7534(22)00006-0. doi: 10.1016/j.annonc.2021.12.015. Epub ahead of print. PMID: 35026413.	BIO recommends the Agency also mention randomized controlled trials using registry data to augment an internal control arm.
Lines 119-122 and 423-425	The draft guidance states, "Before designing and initiating an interventional or non-interventional study	BIO recommends that the Agency consider editing the guidance to recommend that sponsors discuss study designs

SECTION	ISSUE	PROPOSED CHANGE
	using registry data for regulatory decisions, sponsors should consult with the appropriate FDA review division regarding the appropriateness of using a specific registry as a real-world data source." The FDA has an opportunity to communicate to sponsors the criteria FDA will consider to evaluate the appropriateness of registry data as a RWD source. This request is unclear as written and guidance to sponsors who either use or design registries need input from FDA regarding appropriateness.	protocols, analysis plans and data selection before conducting studies (rather than before designing the studies) and specifying the level of information sponsors should share in order to support a robust scientific discussion about the proposed approach. BIO notes that the *Assessing EHR and Medical Claims Data* draft guidance recommends that sponsors meet with FDA before conducting studies. The recommendations for timing of discussions about development programs should be data source agnostic and therefore aligned across all guidance documents. BIO also notes that it may be difficult for FDA to provide robust advice to sponsors about the appropriateness of registries without detailed information about the clinical question, regulatory context and design of proposed studies.
III. DISCUSSION		
A. Using Registry Data to Support Regulatory Decisions		
General comment	Clarity is needed on acceptance of data from a global registry or registry representing specific geographic regions beyond the United States. This is particularly important for rare diseases or outcomes.	BIO recommends that the Agency clarify that global registries and registries representing geographic specific regions may be accepted. As with other aspects of registry studies, this should be discussed with FDA.
Lines 144-154	The draft guidance states, "Registries can have limitations for use in a regulatory context…." As written, FDA presents the limitations mentioned, e.g., disease severity, self-care practices, socioeconomic background, as obstacles to registry use. One way to view the value of registries is that they can be designed to account for and capture data on "real world" factors that impact patient care.	BIO recommends that the Agency consider and provide guidance on how to address these limitations in existing and newly designed registries in order to design registries that are "real world" focused and capture multiple relevant factors that impact outcomes. BIO recommends the Agency consider adding two sentences to this section: **"Sponsor should clearly discuss the registry limitations and impact of such limitations on the study results."**

SECTION	ISSUE	PROPOSED CHANGE
	It is worth mentioning in this section that registries often have incomplete (or omitted) capture of treatment information (e.g., NIH's Surveillance, Epidemiology, and End Results (SEER) registry).	**"In addition, registries often have incomplete capture of treatment information."**
Line 152 - 154	The draft guidance states, "Additional potential limitations of registries involve issues with data heterogeneity…" A strength of registries is to have a sufficient number of patients, especially in rare diseases or rare outcomes. Heterogeneity of the data may be inherent to such registries, which needs to be recognized as an acceptable limitation.	BIO recommends the following edit: "Additional potential limitations of registries involve issues with data heterogeneity (e.g., different clinical characteristics across various populations) and variations in approaches used to address data quality. **A strength of registries is to have a sufficient number of patients, especially in rare diseases or rare outcomes. Heterogeneity of the data may be suitable for certain research questions, such as natural history studies.** "
Line 156 - 161	The draft guidance states, "In general, registries are better suited as a data source for regulatory purposes when sponsors aim to capture objective endpoints, such as death or hospitalization. Subjective endpoints, such as pain, can be collected in a registry, but additional challenges are involved to standardize such measurements. In addition, a registry that is designed to collect data to answer a specific research question can have advantages over an existing registry designed for another purpose, which is subsequently repurposed for that same question." These considerations and limitations are not unique to registries but apply to all RWD data sources. Registries can be extremely valuable and reliable sources of incidence and prevalence data and are commonly used to support regulatory submissions such as orphan drug applications, diversity plans, etc.	BIO recommends the following edit: "In general, registries **(like all clinical data)** are better suited as a data source for regulatory purposes…" BIO also suggests the Agency consider including estimations of disease occurrence as an important use and highlighting in the guidance that both exposures and outcomes should be valid and reliable.

SECTION	ISSUE	PROPOSED CHANGE
Lines 158-159	The draft guidance uses pain measurement as an example of a subjective endpoint with challenges to collect standardized measurements.	BIO recommends the Agency provide additional details regarding standardized measurements for this example. If patient reported pain is captured from all patients using a 0-10 NRS at regular intervals, would this be considered standardized (realizing, of course, the potential for missing data)?
Lines 166 -167	Relevance and reliability are described in the draft guidance as key points for regulatory decision, but these terms are currently not defined in the glossary.	BIO recommends the Agency add a definition for "relevance" and for "reliability" in the glossary.
Line 168	It is not clear if data elements also include registry selection criteria, time period definitions, and key covariates.	BIO recommends that the Agency clarify if data elements also include registry selection criteria, time period definitions, and key covariates.
Line 169	The draft guidance states, "Relevance includes the availability of key data elements (patient characteristics, exposures, outcomes)". In order to accurately estimate the primary estimand, the data source should also capture intercurrent events which could preclude observation of the outcome variable or affect its interpretation (e.g., use of rescue medication, treatment switching or treatment discontinuation).	BIO recommends that the Agency expand this list of key data elements to include "intercurrent events that may preclude observation of the outcome or affect its interpretation."
B. Relevance of Registry Data		
Entire Section	When considering the appropriateness of a registry, potential sample size should be taken into consideration. If it is unlikely that enough patients would be recruited (e.g., due to low exposure during pregnancy for a pregnancy registry), a registry is likely not appropriate.	BIO recommends that the Agency provide additional comments on the challenges and limitations of recruiting and enrolling patients into the registry, and when a registry may need to be discontinued or may be not appropriate. BIO recommends that the Agency provide additional comments on criteria for a registry to be considered representative of the population.

SECTION	ISSUE	PROPOSED CHANGE
		BIO recommends that the Agency provide additional comments on when a comparator group is needed as part of a registry.
Entire Section	There are high quality registries in European countries that accurately follow up patients within their country healthcare systems and can be great sources for registry-based studies.	BIO recommends that the Agency provide its view on using the data of a non-US registry if the data is relevant (good availability of data elements, accuracy follow-up, sufficient subjects and generalizability to target population in the US).
Lines 189-196	The draft guidance states, "The assessment of the data's relevance is context dependent…."	BIO recommends that the Agency consider providing examples and/or specific guidance on how FDA would think through assessment of how well a registry population represents the sponsors target population (i.e., qualitatively, what types of characteristics would the Agency consider and what types of indicators of similarity would the Agency look for?)
Line 194-196	The draft guidance states, "…patients who remain enrolled in a registry may differ from those who do not remain (e.g., having experienced an adverse event)…" It is unclear how having an adverse event (other than death) would influence the ability to follow up on a patient in a registry. It is unclear if the assumption is that the patient needs to remain on treatment in order to remain enrolled in a registry. Patients are lost to follow up for many different reasons and inability to account for these characteristics may result in a biased sample.	BIO suggests clarifying how having experienced an adverse event might influence the ability to follow up on an enrollee. Alternately, we suggest citing an example from literature.
Line 198-202	The draft guidance states, "Registries general include data elements that capture information about patient characteristics, treatments received, and health outcome for patients enrolled in the registry. Such information typically includes a unique patient identifier; the date of patient consent to participate in the registry; and baseline characteristics of the patient at that time,	BIO recommends that the Agency consider clarifying that some registries may be exempt from informed consent.

SECTION	ISSUE	PROPOSED CHANGE
	such as demographic factors, comorbidities, medical history, and other information." Some registries may be exempt from consent and provisions should be allowed for these instances. For example, the antiretroviral pregnancy registry is exempted from obtaining informed consent.	
Line 202-258	The draft guidance states, "Sponsors should consider which data elements a registry should have based on their intended use of the registry. The following are non-exhaustive examples of potential data to include in a registry:" The first sentence appears to refer to using an existing registry data as RWD. The second sentence appears to refer to designing a new registry.	BIO recommends that the guidance be separated into each case (using an existing registry vs. designing a new registry).
Lines 205-258	The draft guidance states, "The following are non-exhaustive examples of potential data to include in a registry:…"	BIO recommends that the Agency consider a core set of data elements to include in all registries and align with either existing FDA data standards and/or USCDI data elements for digital health data. Indeed, the disease/condition-specific registries will have their nuances, but FDA should provide guidance on a core set of demographic, clinical, treatment, health related outcome, and other patient reported data elements for each registry.
Lines 207 -210	The collection of dates (e.g., date of birth) might be limited to year of birth, similarly for other dates, in certain countries.	BIO recommends the Agency mention that the collection of demographics might be restricted by country-specific regulations.
Lines 209-210	The draft guidance states, "Patient demographic factors, including date of birth, gender, race and	BIO suggests the Agency consider adding geography to this bullet, i.e. state/region if in the United States, country if Global.

SECTION	ISSUE	PROPOSED CHANGE
	ethnicity, height, weight, smoking status, alcohol use, and recreational drug use"	
Lines 212-215	The draft guidance states, "Primary diagnosis of interest, including date of diagnosis, test name and result…" For clarity, the guidance should state "diagnostic" test name, to assure accuracy of which diagnostic was used.	BIO recommends the following edit: "Primary diagnosis of interest, including date of diagnosis, **diagnostic** test name and result…"
Lines 235-236	The draft guidance states, "Specific clinical events (e.g., heart attack, stroke, hospitalization, death) of interest and date of occurrence"	BIO suggests adding the following phrase to the bullet point: "specific clinical events (e.g., heart attack, stroke, hospitalization, death) **or other AEs** of interest and date of occurrence"
Line 241	The draft guidance states: "Changes in patient management and date of occurrence" Examples are missing for this item.	BIO recommends providing examples for "Changes in patient management," similar to what is provided for the other two bullets under "Health-related outcomes."
Line 243	The draft guidance states: "Pregnancy-related information" We recommend adding additional considerations to this section.	BIO recommends listing additional potentially important exposure variables including (dose, frequency, duration, and indication) and potential confounders including exposure to known teratogens, smoking behavior, comorbidities, etc. to this section.
C. Reliability of Registry Data		
Entire Section	As drafted, collection of patient-reported outcomes data in the context of a registry to generate information on burden of disease, current unmet treatment needs, and signs/symptoms & impacts experienced by patients to inform medical product development is only mentioned in section C (Reliability of Registry Data).	BIO recommends that it would be helpful if this topic was discussed in preceding sections (e.g., those outlining the relevance of registry data and potential data to include in a registry) as well.
Lines 262-266	The draft guidance states, "When considering using an existing registry…"	BIO recommends that the Agency consider partnerships and collaboration with existing disease/condition-specific registries

SECTION	ISSUE	PROPOSED CHANGE
		that have experience with registry operations. Included in registry operations should be clear guidance on how patients have input into not only the design of registries but access to data used for regulatory decision making.
Line 274-276	The draft guidance states, "Data collection, curation, management, and storage, including processes in place to help ensure that data within a registry can be confirmed by source data (as applicable) for that registry." Source data can vary based on data collection methodology. For example, if a patient reported outcome or healthcare professional reported outcome is entered directly into an electronic data collection system, the source data would be difficult to confirm (i.e., no medical chart notes).	BIO recommends that the Agency clarify what is required for "confirming" source data and consider situations when it may not be possible.
Lines 285 - 287	While there is a cross reference to footnote 12, clarification on the use of foreign registries might also be included directly in this guidance.	BIO recommends that the Agency clarify expectations for the use of registries outside of US in terms of conformance with 21 CFR part 11.
Lines 296-297	The draft guidance states, "Factors that FDA considers when assessing the reliability of registry data include how the data were collected (data accrual)."	BIO recommends that the Agency elaborate on what characteristics of data accrual/collection need to be evaluated to ensure reliability of a registry.
Lines 301-305	The draft guidance refers to the recommendations in the 2009 FDA's guidance; however, the 2009 guidance does not cover capturing PROs in the real world setting.	BIO recommends that the Agency consider updating the FDA 2009 draft guidance for industry entitled, "Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Changes" to reflect the nuances of PRO capture in the real-world setting.
D. Considerations When Linking a Registry to Another Registry or Another Data System		

SECTION	ISSUE	PROPOSED CHANGE
Entire Section		BIO recommends that the Agency consider addressing ways to link and/or mine data through artificial intelligence or machine learning.
Line 289	The draft guidance states, "Sponsors also should ensure that a registry adheres to applicable jurisdictional human subject protection requirements, including protecting the privacy of patient health information...."	BIO recommends that the Agency highlight that as part of the due diligence process, Sponsors should ensure that registry owners/PIs understand applicable human subject protection requirements. The specific requirement should be enforced by an ethics committee.
Line 299	The draft guidance states, "Sponsors should address whether the registry has privacy and security controls in place to ensure that the confidentiality and security of data are preserved." As written, the role of "Sponsor" is unclear.	BIO recommends that differentiation should be made for Sponsors designing and implementing de novo registries as compared with those embedding studies within existing registry systems/databases (i.e., secondary use of data NIS). As for the second situation, additional guidance would be helpful to understand how to ensure privacy and security controls are in place at the time of data collection.
Line 315	The draft guidance states, "Sponsors are encouraged to use common data elements…."	BIO recommends the Agency provide examples of what they consider common data elements. In the case of secondary use of data, this guidance may not apply. It would be helpful to differentiate de novo work from secondary use of data studies and corresponding requirements/suggestions.
Lines 320-322	The draft guidance states, "Appropriate policies and procedures should be in place to support the reliability of the registry data, including prespecifying data validation rules for queries and edit checks of registry data, as well as validating the electronic systems used to collect registry data." It is unclear what FDA may consider appropriate policies and procedures.	BIO recommends that the Agency provide details on the level of expectation for "appropriate policies and procedures," and if they are expected to be documented and shared. BIO also recommends that the Agency consider removing the term "validating electronic systems" or clarify expectations, particularly when the MAH is not the owner of the data.

SECTION	ISSUE	PROPOSED CHANGE
	The expectations for validating electronic health care systems are unclear, and may not be possible by the MAH. This may not always be possible (e.g., validating EMR data system from a hospital system).	
Lines 355-360	The draft guidance states, "Indicators of data consistency, accuracy, and completeness should be assessed periodically, with the frequency dependent on the purposes of the registry data (e.g., for the sole purpose of facilitating recruitment in a randomized controlled trial versus using the registry data in an interventional or non-interventional study analysis)."	To promote uniformity into how these important data quality components are incorporated into registries used for regulatory-decision making, BIO recommends that FDA consider making these required elements and provide clear timelines for when and how registries are updated, refreshed, and reconfigured, particularly those registries that use electronic medical record data. EMR systems frequently update and upgrade (https://www.nature.com/articles/d41586-019-02876-y), thus FDA should collaborate with registry sponsors to help determine criteria for data quality and specific time periods for registry revaluation. BIO also recommends that the Agency provide examples of metrics considered potentially appropriate for evaluating the degree of consistency and accuracy, as these may not be as straightforward to assess as a concept such as data completeness.
Line 367	Data linkages may be deterministic or probabilistic.	BIO recommends that the Agency consider providing more detail on approaches for linking patients between data sources and advantages/limitations. For example, is probabilistic linkage allowable if the false negative and false positive rates are satisfactory? Consider aligning the recommendations with the linkage section in the RWD guidance for EHRs and medical claims.

SECTION	ISSUE	PROPOSED CHANGE
Line 370	Consider providing examples or suggestions of when it is more appropriate to collect continuous vs. intermittent data.	BIO recommends that the Agency consider providing examples or suggestions of when is more appropriate to collect continuous vs intermittent data.
Lines 392-394	The draft guidance states, "Documentation of the process sponsors used to validate the transfer of data…" <u>Data Linkage Partners</u> It is described that documentation of the process that sponsors use to validate the transfer of data should be available to the FDA for review during a sponsor inspection. Sometimes sponsors will partner with a vendor that is performing the linking, including documentation.	BIO recommends that the Agency consider the ability of these vendors to provide documentation to the sponsors or be accessible for inspection.
Lines 405-406	The draft guidance states, "The data can be accurately matched to patients in the registry and whether linking records between the two (or more) databases can be performed accurately"	Two key methods, deterministic and probabilistic, are used for linkage. BIO encourages the FDA to provide examples of how these two methods were used with a specific focus on how these methods can impact the reliability of the registry data. BIO recommends the Agency consider editing lines 405-406 with the following language to account for potential selection bias even after accurate matching: "The data can be accurately matched to patients in the registry and whether linking records between the two (or more) databases can be performed accurately, **and whether the linked patients are a representative subset of patients to meet the research objectives"**
Lines 408-409	There are often minor variations in the definitions of clinical endpoints. For example, "progression" may be	BIO recommends that the Agency clarify and provide examples.

SECTION	ISSUE	PROPOSED CHANGE
	locally assessed or adjudicated, pain relief may be measured by visual analogue scale after 12 weeks or by categories after 16 weeks, "remission" may be within 24 weeks or at 24 weeks. Often, proxies such as treatment discontinuation are used if the formal status as defined in a clinical study is not available.	
Individual level data	Clarity is needed on the expectation for submitting individual level data. These data may be owned by a third party and not available for submission. Clarity is also needed about the expectations when the data comes from a registry when the MAH is the sole sponsor, when it is a multi-stakeholder sponsored registry, or when data is owned by an external party.	BIO recommends that the Agency clarify expectations for individual level data when the data comes from a registry when the MAH is the sole sponsor, when it is a multi-stakeholder sponsored registry, or when data is owned by an external party. In addition to clarifying expectations, it is important that FDA work with stakeholders to better understand the considerations for providing patient level data when the sponsor is not the owner of a given registry and work together to address the challenges.
Data dictionary	If the MAH is not the owner of the data or the registry is linked to a second data source, please clarify the acceptable level of detail the data dictionary should include. Data sources may have different levels of detail, and the MAH may not have the ability to mandate certain data standards (e.g., National Death Index). In these cases, the MAH can link the data, but cannot guarantee similar data standards.	BIO recommends that the Agency clarify the acceptable level of detail the data dictionary should include and acknowledge the level of detail may vary based on source of registry data.
Data Linking	Coordination with organizations like Office of the National Coordinator (ONC) should be considered to improve registry design to link to additional data sources.	BIO recommends that the Agency coordinate with ONC to help leverage the range of interoperability standards (e.g. FHIR), and interoperability technologies (e.g., API's, SMART on FHIR), and other areas of data integration expertise to provide guidance and direction to registry sponsors to apply these resources to registry design to link to additional data sources.
E. Considerations for Regulatory Review		

SECTION	ISSUE	PROPOSED CHANGE
Line 414	Consider elaborating on how to handle missing data.	BIO recommends that the Agency elaborate on how to handle missing data within this guidance or a new draft guidance or workshop.
Lines 424-425	The Sponsor should meet with the relevant review division prior to conducting a study that will include registry data intended to support a regulatory decision. However, sponsors and FDA review divisions have somewhat limited experience meeting to discuss proposed RWE study plans (intended to support new approvals or label changes).	BIO recommends that the Agency provide additional details (process, suggested timing/timelines, anticipated role of RWE subcommittee, documents expected at each stage, e.g., protocol/SAP) regarding FDA expectations for sponsor interactions with the review divisions. BIO also recommends that the Agency consider including experts from the FDA RWE subcommittee and/or Office of Biostats and Epidemiology in the review for methodological aspects along with the FDA disease-specific Divisions.
Line 435	The draft guidance states that "All essential elements of a registry study's design, analysis, and conduct should be predefined," When using pre-existing registries, certain elements may already be known due to previous analyses, publications, etc. The appropriateness of using such registries is then questionable. This section should also contain guidance on the elements that need to be discussed when using a pre-existent registry, e.g., in the context of an external control arm for a trial to be conducted. This reference may provide some useful examples of the types of safeguards that may help ensure the rigor of registry studies. Lessons Learned Using Real-World Data to Emulate Randomized Trials: A Case Study of Treatment Effectiveness for Newly Diagnosed Immune Thrombocytopenia - McGrath - 2021 - Clinical Pharmacology & Therapeutics - Wiley Online Library	When using pre-existing registries, certain elements may already be known due to previous analyses, publications, etc. The appropriateness of using such registries is then questionable. BIO recommends that this section should also contain guidance of the elements that need to be discussed when using a pre-existent registry, e.g., in the context of an external control arm for a trial to be conducted. We also ask FDA to provide recommendations on procedures and processes (e.g., documentation) that may help ensure studies using pre-existing databases meet regulatory expectations. BIO also recommends that the Agency consider clarifying how to deal with pre-specification when using pre-existing registries.

SECTION	ISSUE	PROPOSED CHANGE
Line 436	Consider defining "essential" elements of a registry in the document.	BIO recommends that the Agency define "essential" elements of a registry in the document.
Lines 436-438	The draft guidance states, "All essential elements of a registry study's design, analysis and conduct should be predefined…" <u>Study Elements in Protocol</u> It is noted that for each study element, the protocol should describe how that element will be ascertained from the selected RWD source. However, the term "study element" is not defined and may be interpreted differently by various stakeholders.	BIO recommends that the term "study element" be further defined so that the expectations from the Agency for the protocol are more clear.
Line 440	Sponsors will not always have access to this data.	BIO recommends that the Agency clarify if and when access to patient-level data is mandatory.
Line 440-445	The draft guidance states, "Sponsors seeking to use registry data to support a product's effectiveness and safety in a marketing application should ensure that patient-level data are provided to FDA in accordance with applicable legal and regulatory requirements. If the registry data are owned and controlled by third parties, sponsors should have agreements in place with those parties to ensure that all relevant patient-level data can be provided to FDA and that source records necessary to verify the RWD are made available for inspection as applicable." In reference to providing individual level data from a third party, we agree that the outlined approach is preferred, but this is not always possible. We suggest the recommendation be revised to include a provision that the MAH will facilitate FDA interaction and/or further data analyses with the data owner. For example,	BIO recommends that the Agency consider providing more flexibility in providing individual level data from a third party as this will be challenging in some situations. BIO recommends that the Agency consider revising this recommendation and providing a provision that the MAH will facilitate FDA interaction and further data analyses/access via the data owner. BIO recommends that the Agency highlight that data that cannot be provided to FDA due to privacy and/or IP restrictions be available upon inspection. We recommend that the guidance also state that although it is the gold standard to provide FDA all relevant patient-level data in accordance with applicable legal and regulatory requirements, it is not always feasible to do so due to certain confidentiality and patient privacy laws. Under this circumstance, the final guidance should recommend that sponsors discuss alternative approaches with the Agency in advance of conducting the study to replicate RWE results and/or conduct sensitivity analyses. Transparent

SECTION	ISSUE	PROPOSED CHANGE
	companies like HealthCore and IQVIA have regulatory divisions that do such work funded by the MAH.	communication from the sponsors of how data were collected, curated, and analyzed will be is essential to promote the quality of RWD in a regulatory application. We also recommend incorporating the language on line 181 - 183 from FDA's fourth draft guidance on considerations for RWD/RWE that encourages a discussion of access to patient-level data: "**In the early stages of designing a study intended for use in a marketing application, sponsors should discuss with the relevant review division the expectations regarding access to RWD for their development program.** Sponsors seeking to use registry data to support a product's effectiveness and safety in a marketing application should ensure that patient-level data are provided to FDA in accordance with applicable legal and regulatory requirements. If the registry data are owned and controlled by third parties, sponsors should have agreements in place with those parties to ensure that all relevant patient-level data can be provided to FDA and that source records necessary to verify the RWD are made available for inspection as applicable."
IV. GLOSSARY		
Line 466		As written, data curation appears to be a (possible) step within data transformation. If that is the case, BIO recommends that the Agency clarify within the glossary.