REAL-WORLD EVIDENCE ALLIANCE

info@rwealliance.org
rwealliance.org

January 24, 2022

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

Re: Real-World Data: Assessing Electronic Health Records and Medical Claims Data To Support Regulatory Decision-Making for Drug and Biological Products; Draft Guidance for Industry; Availability (Docket No. FDA-2020-D-2307)

To the Food and Drug Administration:

The RWE Alliance appreciates the opportunity to comment on the draft guidance entitled "Real-World Data: Assessing Electronic Health Records and Medical Claims Data to Support Regulatory Decision-Making for Drug and Biological Products" (referred to here as the "Draft Guidance").[1] We are a coalition of real-world data ("RWD") and analytics organizations with a common interest in harnessing the power of real-world evidence ("RWE") to inform regulatory decision-making to improve patients' lives. Our members have deep knowledge and experience working with healthcare data across disease areas and patient populations, and we aim to bring these collective insights to bear in support of RWE policies.[2]

The RWE Alliance envisions a future in which data generated in everyday clinical practice and everyday life through electronic health records ("EHRs"), administrative claims and billing data, product and disease registries, personal devices, wearables, and health applications will be used to generate evidence that complements clinical trial data to inform regulatory decisions. To achieve this goal, the RWE Alliance advocates for policies that will (1) advance FDA's RWE Framework, (2) encourage the use of RWE to better understand treatment effects in underrepresented populations, (3) enhance opportunities for RWE organizations to consult with FDA, and (4) increase communication on the generation and use of RWE.[3]

Section I of this letter provides an overview of our comments, including suggestions regarding FDA's RWE Program. Section II discusses our general comments regarding the Draft Guidance and FDA's plans to operationalize its recommendations. Section III covers our comments on specific sections of the Draft Guidance. Finally, the appendix sets out our additional comments regarding specific passages of the Draft Guidance.

---

[1] 86 Fed. Reg. 54219 (Sept. 30, 2021).
[2] For information about our members, please see our website, https://rwealliance.org/who-we-are/.
[3] Additional information about what we believe is available on our website, https://rwealliance.org/what-we-believe/.

I. Overview of the RWE Alliance's Comments

We applaud FDA for issuing the Draft Guidance as part of the Agency's RWE Program, consistent with its mandate under the 21st Century Cures Act. Our members appreciate FDA's efforts to provide guidance with recommendations referencing multiple therapeutic areas to help ensure regulatory submissions containing RWD and RWE meet regulatory expectations. We believe RWE has the potential to inform regulatory decisions across therapeutic areas and can improve the lives of a wide range of patients.

We agree with the Draft Guidance's general approach of encouraging the pre-definition of essential elements in protocols and statistical analysis plans ("SAPs"). This approach enhances confidence in RWE study results by reducing the risk of a results-driven selection of study parameters.[4] In addition, integrating transparent, auditable, reproducible, and scientifically valid RWE into regulatory decision-making will facilitate more efficient drug development and enhance understanding of product safety and effectiveness. We appreciate that the Draft Guidance identifies these and other key elements that are necessary components of RWE studies in order to support FDA decision-making.

Our primary focus in the comments is facilitating the successful implementation of FDA's RWE Program. A key factor for this implementation is achieving alignment among stakeholders who work with RWD, as well as among Agency staff, on the regulatory expectations for submissions containing RWD/E. With this goal in mind, we highlight the following points for the Agency's consideration regarding the RWE Program's implementation through the Draft Guidance and otherwise:

- **Increased communication about concrete use cases with stakeholders outside of FDA.** We recommend that FDA incorporate, in the Draft Guidance or separate guidance, additional hypothetical examples of acceptable approaches to resolving issues discussed in the Draft Guidance, e.g., for validating a common outcome of interest and applying quality standards. In particular, it would be helpful to see examples of approaches that meet, and fail to meet, FDA's expectations.

  More generally, we recommend that FDA continue to communicate with external stakeholders about examples in which RWE has been used by FDA in regulatory decision-making. While we acknowledge that FDA plans to publish learnings about the use of RWE in regulatory submissions as part of the PDUFA VII program, "Advancing Real-World Evidence for Use in Regulatory Decision-Making," we believe providing additional examples in the shorter term would be useful. The Agency's press releases highlighting the use of RWE in approval decisions have been helpful, as was the Oncology Center of Excellence's

[4] Orsini LS, Berger M, Crown W, et al. Improving Transparency to Build Trust in Real-World Secondary Data Studies for Hypothesis Testing-Why, What, and How: Recommendations and a Road Map from the Real-World Evidence Transparency Initiative. Value Health. 2020;23(9):1128-1136. doi:10.1016/j.jval.2020.04.002.

landscape analysis of oncology drug submissions that have included RWD.[5] We have appreciated FDA's efforts to share this information. Additional methods of communicating new information might include a dedicated section of the review memoranda published by Agency personnel regarding an approval decision, a centralized and public RWE Program "dashboard" on FDA's website, and/or a periodic Agency report that summarizes actual use cases or aggregates the Agency's observations about RWE studies that have not met regulatory expectations, to the extent permissible under FDA's information disclosure laws.[6] In addition, we would welcome the Agency highlighting its ongoing work in the RWE Program—e.g., pilot projects and other collaborative efforts in which the Agency has been involved—in a public forum, such as on a centralized dashboard.

- **Alignment across FDA offices and staff.** We also appreciate the cooperative effort involved in the co-publication of a guidance by the Center for Drug Evaluation and Research, the Center for Biologics Evaluation and Research, and the Oncology Center for Excellence. A critical piece of the RWE Program's implementation will be ensuring alignment on the Agency's recommendations across internal stakeholders so the Agency's policies and expectations are implemented fully and consistently across review divisions and during application reviews. Whether in the final guidance or other publicly available materials, we would welcome insights into FDA's efforts to ensure Agency staff with key knowledge about RWD/E policies and expectations are able to share information and engage with staff who are newer to the area. The Agency could consider sharing insights into this process through the final guidance, other policies/procedures, or on the centralized and public dashboard for the RWE Program described in the previous bullet point.

- **Channels for seeking FDA meetings outside of an individual sponsor's submission.** The Draft Guidance recommends meeting with the Agency on a number of issues. We encourage FDA to include specific recommendations on when and how an RWD/E organization should request a meeting with the Agency and the best process for doing so. We emphasize that these meetings would be to address issues outside of an individual sponsor's submission—for example, an RWD/E organization may wish to meet with FDA to address a single dataset that has the potential to be used across multiple submissions by multiple sponsors. A meeting could be helpful in this context when a dataset will inform FDA's regulatory decision-making in a therapeutic area or when an RWD/E organization seeks clarification about applying the Draft Guidance's recommendations in a particular use case. We would appreciate the Agency's views on the process for requesting such meetings, including the materials that

[5] Rivera D, Lee JJ, Royce ME, Kluetz PG. FDA Oncology Center of Excellence Landscape Analysis of Real-World Data Submissions for Oncology Drugs. Journal of Clinical Oncology 39, no. 15_suppl. DOI: 10.1200/JCO.2021.39.15_suppl.e18787.

[6] *See*, *e.g.*, FDA, *Examples of Real-World Evidence (RWE) Used in Medical Device Regulatory Decisions*, https://www.fda.gov/media/146258/download.

should be provided in advance to Agency participants and other general expectations.

- **Convening a public workshop to address complex questions.** The Draft Guidance covers complex topics in great detail and will have broad implications for the use of EHR and medical claims data in regulatory decision-making. We suggest that FDA consider holding a public workshop so the Agency, sponsors, RWD/E organizations, and other relevant stakeholders can discuss key topics in greater detail. In particular, we are generally supportive of the data fitness recommendations outlined in the Draft Guidance, but a public workshop dedicated to these expectations and how to define them would be helpful. Other topics for a public workshop could include addressing different durations of continued drug effect as potential sources of bias[7] and issues warranting further consideration across stakeholders, such as what constitutes an acceptable level of missingness, misclassification, and a "high" and a "moderate-to-low" positive predictive value ("PPV").[8] The workshop should also address where the Agency contemplates flexible approaches to the data elements discussed in the Draft Guidance. FDA may also wish to consider establishing a working group with key stakeholders to develop a detailed framework, informed by this guidance, that can be applied for data assessment.

- **Improving the Draft Guidance's accessibility.** The Draft Guidance covers a broad range of topics and recommendations in great detail. This specificity is essential, and we appreciate FDA's efforts to provide this guidance. We suggest the Agency consider enhancing the Draft Guidance's accessibility to stakeholders by outlining key, overarching principles at a high level in one section of the final guidance or in an appendix. Taking this step will be especially helpful as FDA publishes additional guidance documents on RWD/E topics and stakeholders work to understand how the various recommendations interrelate.

We recognize these steps require the use of the Agency's limited time and resources, but we strongly believe that improving the public accessibility of this information will help advance this emerging regulatory area and facilitate the development of new treatment options for patients.

II. General Comments on the Draft Guidance

A. We encourage FDA to acknowledge explicitly in the final guidance that the Agency will consider the specific circumstances of an RWE study, including the clinical and regulatory contexts, when evaluating the suitability of using data from a particular source for a specific regulatory purpose

We recognize the Draft Guidance is non-binding and generally reflects the Agency's current thinking on ideal best practices for RWE studies using EHR and/or medical claims data. We suggest updating the Draft Guidance to state more clearly that

[7] *See* Draft Guidance at Lines 637–39.
[8] *See id.* at Lines 857–59.

different approaches can be appropriate in various cases, including when the Draft Guidance's recommendations may not be feasible, practical, and/or relevant under the circumstances of a specific study. We also encourage FDA to clarify that the focus in designing an RWE study should be on addressing the concerns and limitations identified in the Draft Guidance to the extent possible, as well as explaining why any residual issues do not undermine the study's overall validity. In various circumstances, RWE studies may provide value for regulatory decision-making that outweighs concerns associated with confounding factors or bias, as long as these concerns are managed appropriately.

We understand this to be FDA's intent and believe it should be stated more explicitly in the Draft Guidance by indicating that the Agency will consider the practicality and relative importance of various recommendations in the context of a specific RWE study. All study designs, including randomized trial designs, will have at least some impact from confounders or bias on interpretation of the results. Recent RWD/E studies used in regulatory decision-making have demonstrated the potential for observational studies to support regulatory decision-making and contribute to filling evidence gaps, even while taking a variety of approaches to issues discussed in the Draft Guidance. We believe this experience shows that it is possible for a study to provide supportive evidence by using different methods. We agree with FDA's flexibility in accepting these RWD/E in regulatory submissions and believe that this flexible approach should be reflected in the final guidance.

B. The final guidance would benefit from more discussion of FDA's recommendations for how to successfully use RWD curated from more than one source

RWD can be generated from EHR and claims data sources in multiple ways, such as through (1) direct retrieval from a single source (e.g., a health system's EHR); (2) direct retrieval from a disparate set of sources (e.g., multiple research sites) for a specific project; and (3) compilation of data from a disparate set of sources (e.g., many different data sources such as EHRs, claims datasets, prescription claims data, and other sources of vital records) into a general-purpose, aggregated research repository from which project-specific datasets can be retrieved. Indeed, some aggregated datasets are themselves sourced from multiple aggregated datasets.

We encourage FDA to discuss the broad range of potential data sources in the final guidance, as well as the Agency's current thinking on addressing regulatory challenges with RWE generated from more than one data source. We appreciate that these other use cases can pose additional challenges to ensuring RWD is fit-for-purpose but believe appropriate strategies can remove or mitigate many issues related to confounders and bias. As an example, in certain circumstances, the use of well-qualified, auditable data abstractors could harmonize concepts from different datasets (e.g., never smoker vs. non-smoker) in the creation of an aggregated database. The final guidance could highlight this example and others to explain how RWE can be derived from a well-curated, general-purpose dataset in a manner that mitigates risk of bias, confounding factors, and other potential issues.

C. The final guidance should discuss the distinctions between EHR and medical claims data and differentiate the recommendations for these two data sources

We recommend adding discussion to the guidance that outlines FDA's views on the general differences between EHR and medical claims data, including particular considerations associated with each as data sources. A section addressing this topic at a high level could help stakeholders better understand FDA's views on distinctions between the two data sources. Many recommendations in the Draft Guidance refer to both EHR and medical claims data, so we also encourage FDA to differentiate the Draft Guidance's considerations and recommendations, to the extent any are more relevant for either EHR or medical claims data.[9]

D. References to "electronic health care data" should be clarified

The Draft Guidance uses the term "electronic health care data" in several places, and this phrase could be interpreted broadly as covering both EHR and non-EHR data or as referring specifically to EHR data. These different potential interpretations could cause confusion about the applicability of the recommendations in the Draft Guidance. We recommend revising to clarify the use of this term.

We also suggest identifying any other data sources that FDA views as "electronic health care data" and to explain considerations for using those data. For example, the Draft Guidance recommends using patient data and surveys for validation purposes but does not specifically address regulatory considerations for their use.

E. We encourage FDA to identify additional resources that can serve as a useful supplement to the Draft Guidance's recommendations

We acknowledge and appreciate that FDA is taking a flexible approach to RWD to support regulatory decision-making and is not recommending any specific checklists associated with the recommendations in this Draft Guidance. We suggest that FDA consider referencing peer-reviewed publications or white papers outlining best practices to help sponsors meet the Draft Guidance's recommendations. FDA should not endorse one publication over another, but we do recommend identifying examples of peer-reviewed publications that provide visualizations, templates, or frameworks for creating reliable and relevant RWD. For example, we suggest identifying publications or white papers with readily accepted standards or methods of necessary completeness and how to show it to enhance the Draft Guidance's recommendations.[10]

[9] As an example, section V.C.2 of the Draft Guidance ("Ascertainment of Exposure: Data Source") is more relevant for medical claims data than for EHR data. This section would benefit from a direct statement to this effect.

[10] *See* Draft Guidance at Lines 196–99.

III. Comments on Specific Sections of Draft Guidance

The following subsections provide our comments on specific sections of the Draft Guidance. For ease of reference, the headings for each subsection correspond to the titles used in the Draft Guidance.

A. General Considerations (Section III of the Draft Guidance)

As noted above, we agree with the recommendation for sponsors to submit relevant documentation to FDA before a study is conducted.[11] Standardization of study documents can be beneficial, but we recommend the Agency clarify that this submission may not always take the form of a protocol and an SAP. In some cases, for example, a protocol may contain the requisite level of detail so an SAP would be unnecessary. Submissions could also include alternative documents providing sufficient detail on the design, analysis, and operational definitions of variables. For example, sponsors could submit a detailed protocol and appendix or a detailed protocol and implementation document (e.g., the STaRT-RWE template[12]). We encourage FDA to adopt a flexible approach in these recommendations to account for the different circumstances of specific RWE studies.

We suggest carrying these points throughout the Draft Guidance, where appropriate. As one example, the Draft Guidance says, "The protocol should include a detailed description of methods for determining how inclusion and exclusion criteria . . . will be implemented to identify appropriate patients meeting these criteria from the data source."[13] Information on the completeness of data for implementing inclusion/exclusion criteria may not always be available when the protocol is finalized, so the recommendation as written may not be feasible when unstructured EHR data are used for evaluating these criteria, though the information could be noted in the SAP. We recommend revising this text to say, "protocol *and/or statistical analysis plan*." More generally, where the Draft Guidance recommends documenting an assessment of a particular topic but does not specify which study materials should capture that information, we encourage the Agency to allow for flexibility in practice. This approach is appropriate given the different study-specific circumstances that could arise and affect where information is documented.

B. Enrollment and Comprehensive Capture of Care (Section IV.B.1 of the Draft Guidance)

We recommend introducing distinct terms to represent a patient's time in a health system and a patient's coverage by an insurance program. The Draft Guidance recommends addressing "continuity of coverage (enrollment and disenrollment) . . . when using EHR and medical claims data sources, given that patients often enroll and

[11] *See id.* at Lines 97–99.
[12] Wang SV, Pinheiro S, Hua W, et al. STaRT-RWE: structured template for planning and reporting on the implementation of real world evidence studies. BMJ. 2021;372:m4856. Published 2021 Jan 12. doi:10.1136/bmj.m4856.
[13] Draft Guidance at Lines 543–44.

disenroll when they experience changes in employment or other life circumstances."[14] The excerpted passage appears to be referring to a limitation of medical claims data, as disenrollment from an insurance plan would introduce the described "discontinuity." The Draft Guidance defines "continuity of coverage," however, as the "period of time over which a patient is enrolled *in a health care system* and during which any medical service or drug prescription would be captured in that *health care system's electronic record system*."[15]

To help distinguish these two concepts in the Draft Guidance, we suggest using "continuity of care" when referring to a patient's enrollment in, or regular use of, a particular health care system. This term should be distinct from "continuity of coverage," which we view as more applicable for medical claims data.[16] In contrast to "continuity of coverage," an issue with "continuity of care" would arise if a patient seeks care outside of a system included in an EHR network (e.g., a specialist in a different health system) and that information is unobservable to the selected real-world dataset. We also suggest the Agency acknowledge that even when there is no clear record of enrollment or disenrollment from a health care system, proxy measures of continuity (such as lengthy absence from typically regular care) can be used. Such proxy measures can be described in submitted protocols, SAPs, and related documents.

More generally, we suggest incorporating the concept of observability into the Draft Guidance and focusing less on continuity. Observability is a key factor in evaluating whether a dataset is likely to reflect relevant patient information (e.g., exposures and outcomes) during a particular time interval. If a patient's observability status is unknown, then one would not know if the absence of a health care event (e.g., a stroke) means the event did not happen—or that it happened but was not recorded. Thus, knowing the time intervals over which a patient is observable is important to applying appropriate analytical methods to deal with periods of unobservability.

C. Data Linkage and Synthesis[17] (Section IV.B.2 of the Draft Guidance)

We appreciate the Draft Guidance's focus on data linkage and synthesis. We agree that linking data sources can provide a more complete picture of a patient's health, and we appreciate FDA acknowledging the benefits of different linkage approaches.[18] In our view, a well-designed and well-executed methodology should be the most important consideration for determining an effective approach to data linkage.[19] We suggest the Agency underscore this overarching principle in the final guidance.

---

[14] *Id.* at Lines 203–05.
[15] *Id.* at Lines 1287–89 (emphasis added).
[16] We also recommend adding "continuity of care" as a defined term in the glossary.
[17] We have used the Draft Guidance's heading for this subsection for ease of reference. Our suggestion, however, would be to revise this heading to "Data Aggregation From Multiple Sources" to capture the subject matter being discussed.
[18] Although data linkage can be a good tool for understanding an individual patient's information, it can have a significant impact on patient attrition due to low expected linkage rates across many datasets.
[19] Doidge JC, Harron K. Demystifying probabilistic linkage: Common myths and misconceptions. Int J Popul Data Sci. 2018;3(1):410. doi:10.23889/ijpds.v3i1.410.

We also encourage FDA to clarify the discussion of heterogeneity[20] to acknowledge the benefits of heterogeneity in patient population characteristics and clinical practices. For instance, heterogeneity in population characteristics can provide additional insight into diverse and underrepresented groups, and these insights can help advance the important public health objective of addressing disparities in healthcare access, treatment, and outcomes. We agree with the importance of accounting for how such heterogeneity may or may not yield heterogeneous treatment effects, and recommendations from the Agency to account for this possibility (e.g., through subgroup analyses) would be beneficial. Separately, recommendations for addressing heterogeneity in coding practices would also be appreciated. Important considerations for this issue include feasibility, ensuring the consistency of clinical concepts, and the benefits of bringing more sites—and thus more patient experience—into a study.

We also seek clarification on the following points:

- The Draft Guidance recommends demonstrating whether and how data from different sources "can be obtained and integrated with acceptable quality . . . ."[21] We suggest clarifying the Agency's expectations for "acceptable quality" or to provide an example of an appropriate approach.

- The Draft Guidance also recommends documenting approaches taken to "address issues that cannot be fully rectified by curation."[22] We suggest the Agency provide more context and an example to help clarify what "issues" are referenced.

D. Distributed Data Networks and Common Data Models (Section IV.B.3 of the Draft Guidance)

We agree that Common Data Models ("CDMs") can be a useful tool because of their ability to execute a query across disparate data sources, and we appreciate FDA's consideration of various types of data networks and the potential benefits and drawbacks of working with CDMs. We encourage FDA to provide further guidance differentiating between the types of CDMs, their strengths and limitations (such as feasibility depending on the data source), and their implications for an RWE study's validity. We also encourage FDA to clarify how sponsors should apply the considerations discussed in the Draft Guidance when submitting a study for which data must be submitted using a standard specified in FDA's Data Standards Catalog.[23]

E. Validation: General Considerations (Section IV.D of the Draft Guidance)

We agree with FDA about the importance of validating variables of interest in RWE studies and appreciate the Draft Guidance's recommendations on this topic. The

[20] *See* Draft Guidance at Lines 269–72.
[21] *Id.* at Lines 270–71.
[22] *Id.* at Line 286.
[23] *See* FDA, *Draft Guidance for Industry: Data Standards for Drug and Biological Product Submissions Containing Real-World Data* (Oct. 2021) ("Data Standards Draft Guidance"), https://www.fda.gov/media/153341/download. The RWE Alliance will submit additional comments on this topic in a separate response to the Data Standards Draft Guidance.

necessary degree of validation will depend on the individual study question or use case. The final guidance should clearly state that the approaches to outcome validation highlighted in the Draft Guidance are examples and that the actual approach taken in a specific instance will be study-dependent.

For example, the Draft Guidance explains that complete verification of a study variable may not be feasible due to factors such as a very large study population or the lack of reference standard data for all study subjects.[24] We appreciate the Agency's recognition of this point. In these scenarios, other methods can be used to improve the interpretability of data. For example, researchers can utilize analogous validated algorithms, which are validated in a different data source with similar parameters, or in a similarly structured study using the same data. Under this approach, the objective would be to demonstrate similarity to known cases. Additionally, researchers can conduct sensitivity analyses to improve interpretability.

We suggest updating this section to highlight these additional options that may be considered when designing an RWE study. This clarification would be consistent with our suggestion in section II.A of these comments—i.e., to acknowledge explicitly that different approaches can be appropriate in various cases and FDA will evaluate what is suitable under the circumstances. We also suggest that the Agency articulate a risk-based approach for validating study variables, under which the extent to which a data element is validated depends on its importance to the overall study. We believe this approach is consistent with the Draft Guidance's recognition that complete verification of a variable may not be warranted in various cases and with the shift to risk-based approaches in clinical trial data monitoring.

Recommendations for validating composite endpoints consisting of many specific endpoints, such as total congenital malformations, also would be beneficial—either in the final guidance or in future guidances. For example, we suggest clarifying whether FDA recommends separate and high-performing algorithms for each specific birth defect or if a more broad-based algorithm discerning a congenital malformation from the absence of malformations would be acceptable.

We also seek clarification on the following points:

- Footnote 10 of the Draft Guidance states that "complete verification involves assigning an accurate value to the variable of interest for each study subject based on a reference standard of choice." The Draft Guidance then goes on to say that "medical record review can be used in conjunction with a conceptual definition to determine whether a subject meets a critical inclusion criterion or has experienced the outcome event." This footnote appears to imply that abstraction (i.e., medical record review) can qualify as "complete verification" in appropriate circumstances. We agree and recommend clarifying the language in the footnote to confirm. We also welcome the addition of other approaches, such as AI-driven

[24] Draft Guidance at Lines 464–65.

validation in which human validation of a subset of cases is used to train an algorithm that completes the validation.

- We suggest outlining the Agency's recommendations for selecting a reference standard and the recommended approach if there is no suitable reference standard.

IV. Additional Comments

The RWE Alliance offers additional comments and suggested clarifications on the Draft Guidance in the Appendix, which begins after the signature page.

V. Conclusion

The RWE Alliance appreciates the Agency's commitment to advancing the use of RWD and RWE in regulatory decision-making. We commend FDA's careful consideration of the topics covered by this Draft Guidance, and we look forward to the other forthcoming guidances, including the guidance on study designs that is referenced in the Draft Guidance.

Thank you for considering these comments, and please let us know if you have any questions. We would welcome the opportunity to discuss further.

Best regards,

The RWE Alliance

Appendix: Additional Comments

Line Numbers	Comment
51–69	We suggest clarifying which topics listed in lines 54–61 apply to relevance (#1 and #2) and which apply to reliability (#3). Moreover, given that the definitions of "reliability" (line 66) and "relevance" (line 67) are critical, and the terms first appear in lines 51–69, we suggest defining those terms earlier in the Draft Guidance, as that will help with reading comprehension.
146–48	The functionality and configuration of a health care system's EHR system are additional factors that can determine how data are recorded. We suggest that FDA consider noting these points as well in the referenced passage.
148–49	We recommend clarifying in this sentence that EHR data consist of structured and unstructured data and that unstructured data include abstracted data.
152–56	We suggest modifying this passage to reflect a few points regarding the prospective use of EHR data. As an initial matter, we suggest acknowledging explicitly that EHR-based data may be comprehensive when collected prospectively, even if an EHR system may not allow for extensive data capture in other cases. Second, prospective clinical studies proposing to use EHRs would more likely involve a configuration of the EHR system to collect and/or require the collection of certain data elements in a certain format, rather than modification of the EHR system through an add-on module. Finally, the text refers to a "limited ability to add modules to collect extensive additional information." Even if the second point is addressed, we see the issue not as a functional limitation for most EHR systems; instead, we see it as a potential impact on provider EHR workflows at the point of care and the challenges associated with modifying a very large enterprise software system for specific use cases, such as a clinical study.
158–59	We suggest clarifying that the description of historical use can rely on publications about the use of a specific data source to answer a question when such publications are available, as opposed to a *de novo* quantitative assessment of accuracy for study elements such as the inclusion/exclusion criteria, outcome, and covariates.
169–73	We suggest adding that the jurisdiction in which the health care system or insurance program is located can be another factor. For example, the local regulatory environment can also have a significant impact on which data elements providers are required to document (e.g., for purposes of billing or quality oversight) and the manner in

Line Numbers	Comment
	which they do so. These variations may emerge in configuration differences across systems and programs. As another example, terminology in medical documentation can also have significant differences across jurisdictions.
176–85	The recommendations in this passage are most relevant for EHR data sourced from a single health care system or datasets that are smaller and less heterogeneous. But in practice, and as noted in our comments above, many data sources span across health systems and/or multiple payers. While we agree with the first recommendation (that the reason for selecting the particular data sources, even whether there is more than one, should be described), the recommendations for demonstrating relevance may not be practical when merging data from larger data networks in which there may be various preferred treatments, diagnostic methods, formulary decisions, and/or patient coverage. For example, it generally is not possible to identify preferred prescribing practices without working at a particular health system, and formulary decisions and tiering levels can change frequently over time (and, in any case, are not public information). In our members' experience, sourcing data from multiple health care systems or multiple payers can reduce any biases arising from a single-system source, such as bias arising from the particular prescribing or care practices, or the specific formulary, at a single location or network. We recommend revising this discussion to acknowledge the potential benefits of multi-system and multi-payer approaches and to identify any relevant considerations or recommended descriptions from multi-system or multi-payer data. As an example, this section could reflect that data may be received from an aggregator of various sites and/or systems, so it could be difficult to create a "profile" that is unique to each site for larger health care systems and may be infeasible for medical claims.
196–99	We recommend clarifying the parenthetical "(e.g., timing of exposure, timing of outcomes)." It would be helpful to understand, for example, if this refers to the exposure and outcome assessment windows or instead to the hypothesized "time at risk" (including when a beneficial effect could be expected). We would also add that another consideration, in addition to the ones highlighted in the Draft Guidance, is whether outcomes of interest would be seen and recorded at a particular facility (e.g., if particular health issues may be sent to specialists in another network and thus not captured in the relevant EHR system).
212–13	We suggest modifying this discussion to reflect that a data source's comprehensiveness should be informative but not dispositive of whether it captures information that is relevant to the study question. Even if a data source

Line Numbers	Comment
	appears to include all components of care available, that does not mean the information is complete, especially for aspects of a patient's history such as hospitalization.
217–19	We believe the examples discussed here reflect a study design consideration rather than a data capture consideration. If clinicians do not typically order a test at a particular frequency, that is a reflection of care practices rather than data comprehensiveness. We suggest omitting this example or replacing it with an alternative example that better reflects considerations specific to data capture. Another example could be that a patient's labs are done in a lab network outside of those that are recorded in the data at hand.
231–35	We suggest expanding this discussion to clarify how information gaps may be addressed. As revised, the sentence would read in part, ". . . *or the protocol should describe how this information gap will be addressed (e.g., by building additional modules into the EHR system or collecting additional data outside of the EHR system)*." In addition, we suggest the Agency consider expanding "immunizations offered in the workplace" to say, "offered in the workplace, *at pharmacies or public health clinics, or through government immunization programs*." This edit will reflect more completely the range of contexts in which medical products are used that may fall outside of typical EHR and medical claims data capture.
241–46	We suggest FDA modify the last sentence to read, "If these issues are relevant to the study question of interest, the protocol should *explicitly* describe *such limitations and the potential impact on the analysis, including any elements of the study design that address such limitations*." In general, quantitative bias analysis could be useful in determining the bounds of error. For example, if we fail to account for 35% of exposures, we would see bias of x%. If the resulting bias is small, the potential impact on the study may be negligible, and it may not be possible in some cases to address these issues. For example, RWD cannot capture behavior considerations or care-seeking behaviors. We believe the proposed modification acknowledges this possibility while also encouraging additional specification in the protocol as needed.

Line Numbers	Comment
379–84	We understand FDA's affirmative statement that the Agency does not endorse any specific AI technology. We believe it would be helpful, however, for FDA to clarify its recommendations for the use of AI in this particular context. We suggest expanding this section or issuing a separate guidance to address this topic.
392–93	We suggest clarifying that this recommendation specifically addresses the use of AI or other computerized extraction methods, not human abstraction. The revised sentence could read, "Relevant impacts on data quality *from use of AI or other computerized derivation methods* should be documented in the protocol and analysis plan."
397–401	We recommend further characterizing the two types of missing data to enhance reader comprehension. The case of "traditional missing data" could be identified as "data missing at the patient level." The second case, of data not intended to be collected, could be described as "data missing at the source level" or "unrecorded data." We also suggest highlighting the importance of sensitivity analyses in this discussion of missing data, which was a key finding from analysis of the GRACE Checklist.[25]
402–19	We recommend revising this text to say, "the decision not to order the test or a patient's decision to forgo the test *(to the extent this is known)* may have implications[26] . . . The protocol and the statistical analysis plan should be developed and based on an understanding of reasons for the presence and absence of information *(if known)*." These edits would be consistent with the Draft Guidance's discussion of insurance formulary changes,[27] and adding the italicized text here is critical because it is often impossible to determine intent with RWD, particularly for missing data. There can be any number of possible explanations for an order not being placed, as the Draft Guidance acknowledges in the excerpted passage, and the reason may not be captured in the notes or otherwise be possible to understand due to de-identification of the patient data. Moreover, a provider and a patient may not necessarily use or have a single EHR; they may instead use or have several EHRs that are not interoperable. The Fast Healthcare Interoperability Resources standard, which allows patients to download their respective EHRs,

[25] Dreyer NA et al. The GRACE Checklist: A Validated Assessment Tool for High Quality Observational Studies of Comparative Effectiveness . J Managed Care & Specialty Pharmacy 2016; 22(10) 1107-1113.
[26] Schneeweiss S, Rassen JA, Glynn RJ, Myers J, Daniel GW, Singer J, Solomon DH, Kim S, Rothman KJ, Liu J, Avorn J. Supplementing claims data with outpatient laboratory test results to improve confounding adjustment in effectiveness studies of lipid-lowering treatments. BMC Med Res Methodol. 2012 Nov 26;12:180. doi: 10.1186/1471-2288-12-180. PMID: 23181419; PMCID: PMC3533513.
[27] Draft Guidance at Line 533.

Line Numbers	Comment
	served as an important step forward. Even with this progress, however, not all of the notes are downloadable, and this limitation restricts the broad utility of downloaded notes as relevant data may not be available. Our proposed edits would make these recommendations more feasible to implement, given the challenges with determining whether data are intentionally missing or not. These edits would also be consistent with the Draft Guidance's objective of removing subjectivity from the collection and assessment of RWD.[28] We also suggest explaining how FDA evaluates whether a strategy for handling missing data is adequate. This passage mentions data linkage and identifying a proxy for the missing data as two potential options,[29] but the Draft Guidance should clarify how the Agency evaluates whether the analysis is acceptable when key data are missing information for some of the patients.
530–32	We suggest discussing FDA's current thinking, either in this section or future RWE guidance on study design and analysis when using RWD sources, about addressing time periods associated with no (or limited) access to health care services, such as during a pandemic. Although the COVID-19 pandemic is a singular event that is still ongoing, its national and international impact warrants discussion given the potential impact on RWD sources.
569–71	We suggest adding in this section the importance of including the clinical rationale when defining and developing variables in RWD. In particular, we recommend acknowledging that variables do not carry the same weight. A minor confounder will not have the same importance as the study exposure in question or the outcome(s) of interest. Quantitative bias analysis is also a useful tool for evaluating variables in RWD.
683–90	For prospective studies in which the EHR is being purposefully modified to collect data, patient-reported information could also be useful in validation. Patients enrolled in the study could use a patient portal or mobile app, for example, to report medication adherence, and this information in turn could be used for validation. We suggest FDA update this passage to incorporate this additional example.

[28] This comment also applies to the discussion on lines 420-422 of the Draft Guidance, which state that "[a]ssumptions regarding the missing data (e.g., missing at random, missing not at random) underlying the statistical analysis for study end points and important covariates should be supported and the implications of missing data considered."
[29] *See* Draft Guidance at Line 413.

Line Numbers	Comment
733–36	We suggest modifying this passage to reflect that the likelihood of capturing certain outcomes can be study-specific, with potential variation by disease type and the extent to which the patient has ongoing, close encounters with health care professionals that could lead to additional information being documented in medical records. For example, a patient who is pregnant or receiving regular cancer treatment may have more regular and close health care interactions than a patient in a population that is not closely monitored.
747–51	Timeframe is an additional concept of a condition that may be relevant for defining an outcome of interest. We recommend updating this passage to add this point.
820–27	As the Draft Guidance acknowledges, there are occasions where complete verification is not feasible and alternatives to assure sufficiently valid capture may be acceptable given the need for evidence. With this in mind, we recommend clarifying in this passage FDA's expectations for validation with respect to de-identified, commercially available datasets, as well as recommended feasible approaches for providing an appropriate level of validation. We suggest FDA begin this section by noting that the appropriate level of validation may depend on the regulatory use case and the prior experience with a particular data source or outcome variable. We also suggest adding that externally accepted, "gold standard" data sources can be one option for validation in this context. For example, the National Death Index is often considered the gold standard for U.S. mortality data. We suggest revising this passage to add, after "paper format," "*or by using broadly accepted external sources (e.g., mortality data sourced from the National Death Index)*."
836–38	We suggest clarifying these lines to facilitate reader comprehension. In particular, we find it unclear whether the estimated medical record retrieval rate is at the cohort level or individual level. We assume the passage refers to the former, which generally can be estimated (such as by calculating the number with structured data and the number with unstructured data), whereas the latter is difficult to determine for each patient in medical claims data (as opposed to EHR data).
1011–14	This subsection does not specifically discuss social determinants of health, but they can also be important modifiers. It would be helpful to understand if FDA considers geocoding at the ZIP5 or ZIP3 level useful to examine effects within subgroups defined by socioeconomic and other social determinants of health.

Line Numbers	Comment
1016–40	We suggest adding in this section that quantitative bias analysis can also be a helpful approach for various purposes, such as understanding impacts from confounding that may arise from inaccurate or missing data on key confounders. We would also appreciate insight into whether FDA recommends a specific approach (e.g., e-value or array approach).
1122–23	Data quality during data accrual could be supported with change logs, if feasible, to maintain transparency and assurance of study validity. We suggest updating the Draft Guidance to reflect this point.
1156–57	We suggest clarifying FDA's expectations, with illustrative examples, for documentation when changes were made to specific data elements from an original extraction to the final analytic dataset. Our proposed general approach is to provide a high-level description of the data ingestion process.
1159–60	This excerpt speaks to the ability to re-identify unique patients in the source data, but de-identified data typically are encrypted to prevent this re-identification from occurring. We recommend revising this text to state, "De-identification of patient records and *any process that could be used to* re-identify unique patients in original source data without losing traceability (e.g., use of linkage tokens)."
1195–97	We suggest clarifying what FDA sees as relevant considerations for ground truth or a "gold standard," potential discrepancies in the values of a specific variable for a specific patient at a specific time from different data sources, and potential independent adjudication (including appropriate blinding when extracting data to end up with a final analysis dataset). To the extent there are discrepancies in data from different sources, we recommend that these discrepancies be arbitrated using a hierarchical algorithm of "truth," with FDA identifying references that are considered "gold standard" for key data elements (e.g., mortality). We also recommend that FDA encourage sponsors to determine if discrepancies are plausible and thus potentially reflect accurate data (e.g., different lab values may arise from different samples).

Line Numbers	Comment
1251	We recommend defining additional terms used in the Draft Guidance—such as "metadata,"[30] and "data linkage"[31]—in the glossary as well as in the main text.
1300–02	We suggest including abstraction of unstructured data sources in this definition for completeness.
1362–63; 1370–71	The glossary defines "sensitivity" and "specificity" in terms of disease measurement. However, these performance metrics can be applied to other types of variables. Indeed, section IV.D of the Draft Guidance ("Validation: General Considerations") discusses sensitivity and specificity in more general terms. Thus, to be aligned with the Agency's discussion elsewhere in the Draft Guidance, we suggest defining "sensitivity" as "The probability that a classification result will be positive when the subject has the characteristic being measured (e.g., disease)" and "specificity" as "The probability that a classification result will be negative when the subject does not have the characteristic being measured (e.g., disease)."
1378–79	This definition refers to validation conducted "usually according to a reference standard." However, it is not clear how one would validate a definition or measure if there is not a reference standard against which to assess its performance. In this instance, one can perform quantitative bias analysis to assess the impact of various hypothesized extents of measurement error on the interpretation of the results. We recommend incorporating this point into the final guidance, including in the glossary definition of "validation."

[30] *See id.* at Lines 360–62.
[31] *See id.* at Line 250.