January 21, 2022

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

Docket No. FDA-2021-D-0548

RE: FDA Draft Guidance on Real-World Data: Data Standards for Drug and Biological Product Submissions Containing Real-World Data

Janssen Research & Development, LLC (Janssen), one of the Janssen Pharmaceutical Companies of Johnson & Johnson, appreciates the opportunity to provide comments to the U.S. Food and Drug Administration (FDA) Draft Guidance on Real-World Data: Assessing Electronic Health Records and Medical Claims Data to Support Regulatory Decision-Making for Drug and Biological Products.

At Johnson & Johnson, we believe good health is the foundation of vibrant lives, thriving communities and forward progress. That's why for more than 130 years, we have aimed to keep people well at every age and every stage of life. Today, as the world's largest and most broadly based health care company, we are committed to using our reach and size for good. We strive to improve access and affordability, create healthier communities, and put a healthy mind, body, and environment within reach of everyone, everywhere. We are blending our heart, science, and ingenuity to profoundly change the trajectory of health for humanity.

Janssen is a committed leader in the development of real-world evidence (RWE) tools and the use of RWE to improve understanding of medical product safety and effectiveness. We innovatively incorporate RWD onto our development programs and strive to identify opportunities to share learnings across industry. We play key roles in various public-private partnerships, stakeholder groups, and meetings convened to advance the use of RWE in drug development and review (e.g., Duke-Margolis RWE meetings, National Academies of Science, Engineering and Medicine, and the Observational Health Data Sciences and Informatics) and we see value in FDA-Sponsor engagement to advance RWE and utilization for regulatory decision making.

Janssen commends FDA for its commitment to advance RWE by issuing draft guidance focused on the standards for drug and biological product submissions containing real-world data (RWD). The guidance outlines important information regarding RWD management and considerations for sponsors as they conform to data standards currently supported by FDA. Janssen notes that in several locations, the guidance indicates challenges that may be encountered when sponsors attempt to conform to data standards, but the guidance lacks actionable solutions to such challenges. Additionally, the expectations for standards are difficult to interpret and implement without an updated Data Standards Catalog. Janssen recognizes that the draft guidance indicates that additional guidance on the topic may be developed and/or the Data Standards Catalog updated to reflect standards for study data derived from RWD sources. We request that as any additional guidance is developed, FDA consider including suggestions for addressing challenges outlined in the present guidance. It would also be helpful for stakeholders to understand the timeline for FDA to issue additional guidance and/or to update the Data Standards Catalog.

Mechanisms for RWD Mapping and Associated Data Models

Janssen appreciates FDA's efforts to outline possible mechanisms for mapping and transforming RWD. However, in the present draft guidance, FDA's reference to mapping is limited to Study Tabulation Model (SDTM), which may constrain options for data formatting and future data standards to be developed. Given challenges of mapping to SDTM for some RWD sources and types, there may be circumstances when it would be most appropriate for sponsors to map directly using another data standard. In other circumstances, there may not be an existing data standard appropriate for a particular type of RWD. We recognize that in such circumstances, on a case-by-case basis, sponsors may need to first have discussions with FDA to determine the best mapping approach. To mitigate uncertainty and to support the use of the most appropriate approach, we request that FDA indicate in the guidance that other mapping approaches may be appropriate and may be used, pending discussions with FDA. The guidance would also benefit from FDA referencing other data models (e.g., Fast Healthcare Interoperability Resources (FHIR), Sentinel, Observational Medical Outcomes Partnership (OMOP) CDM, Analysis Data Model (ADaM), PCORnet CDM, among others) as potentially feasible models for supporting RWD and associated regulatory submissions, pending discussions with FDA as well as greater detail regarding when data standards may or may not be required and when waivers may be considered. We request that FDA consider including in the guidance, types of submissions which include RWD that would be suitable or not suitable for consideration of waivers for conversion to the SDTM standard, additionally, we also request that FDA reference, under circumstances when a waiver is granted, alternative standards to SDTM that may be acceptable in submissions which include RWD. Finally, there are also probabilistic algorithms such as imaging and natural language processing that may not map well to standard data formatting procedures but represent novel approaches for RWD. We encourage FDA, in partnership with industry, to expand and validate data models that may use these novel approaches and to outline learnings in future guidance and to, when possible, align with other health authorities.

Contribution of RWE Stakeholders

Janssen recognizes that the FDA guidance is written for stakeholders who plan to submit RWD to support regulatory decisions; however, the RWD ecosystem consists of stakeholders who support RWD throughout the lifecycle of the data. Such stakeholders include data generators (e.g., patients, health care facilities, insurers), data aggregators (e.g., companies that aggregate EHRs and medical claims data across multiple data streams), and data consumers (e.g., pharmaceutical companies, research institutions). In some cases, data relating to transformations, provenance of variables and individual data values may require proprietary or protected information from data partners. Additionally, submissions of RWD may often include fields that have been subject to multiple data transformations conducted by stakeholders other than the sponsor. The guidance and the entire FDA guidance series on RWD and RWE would be significantly strengthened if FDA differentiated the various stakeholders and how the stakeholders may work together to support the considerations outlined in the guidance, including acknowledging when an activity may be carried out by a data partner versus the sponsor. It may also be beneficial for FDA to issue guidance describing how to document quality control processes and transformations used to create datasets, including examples of how to document transformations used to create human-abstracted and computationally derived data elements.

Meeting Opportunities to Support Robust RWD Studies

Janssen appreciates FDA's emphasis on engagement between FDA and sponsors. We also note that the PDUFA VII commitment letter emphasizes FDA's commitment to engaging with sponsors on RWD through the Advancing RWE Program. We agree that such meetings hold important value for

both FDA and sponsors as both parties gain experience with RWD and serves as an opportunity to advance the field and use of RWD for regulatory decision making. We do request that the present guidance and the RWD/RWE guidance series include additional detail regarding meeting opportunities, including how sponsors may engage as appropriate (e.g., review divisions in addition to the FDA Real-World Evidence Subcommittee), meeting type and associated timelines, and when sponsors when should engage FDA.

Please also find attached to this letter, additional specific comments pertaining to the draft guidance.

Thank you for your consideration of our comments. We would welcome the opportunity to provide further input on the topics described and/or input on any other FDA RWD work products. For any follow up discussion, please contact Danielle Friend (deconomo@its.jnj.com).

Sincerely,

Najat S. Khan Ph.D

Najat S. Khan, Ph.D.
Chief Data Science Officer, Janssen R&D Data Sciences
Global Head, Janssen R&D Strategy and Operations
Janssen Research & Development, LLC

Karin VAN BAELEN

Digitally signed by Karin VAN BAELEN
DN: c=US, o=JNJ, ou=Subscribers, 0.9.2342.19200300.100.1.1=10011455, cn=Karin VAN BAELEN
Reason: I am certifying this document.
Date: 2022.01.17 09:22:48 +01'00'
Adobe Acrobat DC version: 2015.006.30527

Karin van Baelen, Pharm D
Head, Global Regulatory Affairs
Janssen Research & Development, LLC

JOANNE WALDSTREICHER MD

Digitally signed by JOANNE WALDSTREICHER MD
DN: c=US, o=JNJ, ou=Subscribers, 0.9.2342.19200300.100.1.1=191817, cn=JOANNE WALDSTREICHER MD
Reason: I attest to the accuracy and integrity of this document.
Date: 2022.01.18 07:57:52 -05'00'
Adobe Acrobat version: 11.0.20

Joanne Waldstreicher, MD
Chief Medical Officer
Johnson & Johnson

Section	Issue	Proposed Change
Regulatory Background		
Line 36	In this section FDA defines Real World Data	Janssen suggest including this definition as part of the Glossary.
Lines 56, 58, 110	In this section FDA uses the term "nonclinical" to refer to pre-clinical studies/data but as RWD studies are not clinical studies (but are also not nonclinical depending on the context), there is some ambiguity.	As RWD is not considered "nonclinical" data, Janssen proposes ensuring there is a clear distinction between "nonclinical" and RWD in the guidance and the referenced FDA Data Standards Catalog.
Applying Currently Supported Data Standards to Study Data Derived from Real-World Data Sources		
Lines 91-93	In this section, FDA discusses aspects of health care data that can affect the overall quality of the data.	Janssen requests the following edits: "… (4) the many aspects of health care data that can affect the overall quality of the data, including business processes and database structure, **lack of criteria or different versions used for disease/performance evaluation**, inconsistent vocabularies and coding systems, and de-identification methodologies used to protect patient data when shared...". Additionally, the guidance would be well served with more detail on data standards related to data linkage and patient record linkage across RWD data sources. Janssen suggests the guidance should address the need for standard language in consent forms to provision patient authorization for additional data linkage activities to available real world data sources. Additional regulatory recommendations should address the need for privacy preserving record linkage in observational RWD studies which have a waiver of informed consent.
Line 97	In this section FDA references "data curation" and indicates that "adequate process should be in place to increase confidence in the resultant data."	For clarity, Janssen requests the following edits: "During data ~~**curation**~~ **capture** and data transformation, adequate processes should be in place to **maintain or to ensure** ~~**increase**~~ confidence in the resultant data".

Line 99	In this section FDA indicates that documentation of data curation and transformation may include but are not limited to electronic documentation (i.e., metadata-driven audit trails, quality control procedures, etc.) of data additions, deletions, or alterations from the source data system to the final study analytic data set(s).	Janssen requests that FDA reference "Data Integrity and Compliance with Drug CGMP Questions and Answers Guidance for Industry"[1] to clarify FDA's expectations for "metadata-driven audit trails".
Lines 115-116	In this section FDA mentions data standards approved by FDA but the section does not include reference to such standards.	To avoid interpretation issues, Janssen requests FDA define which standards FDA is referring to in this section. It would also be helpful for FDA to provide examples of such data standards and related transformations, conversations, and mappings that should be considered. In addition, Janssen proposes the following edit: "When seeking to conform RWD to data standards supported by FDA, sponsors should consider ~~**the relevant**~~ data transformations, conversions or mappings that may be needed to produce study datasets in the required format in an applicable drug submissions."
Lines 121-123	In this section, FDA indicates that "Sponsors should describe these approaches, including in the protocol, data management plan, and/or final study reports."	Janssen suggests that FDA considers adding statistical analysis plan (SAP) along with protocol to describe these transforming approaches in the case that the transformation is originated from appropriate statistical analysis, representing a pre-planned consideration. However, a sponsor should not need to provide the details in SAP if the data transformation is not relevant to analysis purpose. To this end, Janssen also requests the following edit: "Sponsors should describe these approaches, including in the protocol, data management plan, **statistical analysis plan**, and/or final study reports, **as applicable**"
Lines 129-131	In this section FDA references study data derived from RWD and how these can be transformed and submitted to FDA in applicable submissions.	For clarity, Janssen requests the following edits:

[1] FDA Guidance for Industry: Data Integrity and Compliance with Drug CGMP Questions and Answers.

		"With adequate documentation of the conformance methods used and their rationale, study data derived from RWD ~~**can be transformed to SDTM datasets and submitted to FDA in an applicable drug submission**~~ **should be submitted if the RWE derived from the RWD is primary evidence in the submission but does not need to be submitted if the RWE is only supportive evidence. What is considered primary evidence should be discussed with the agency**."
Lines 136-137	In this section FDA references data domains and data standards stating: "...there is wide divergence in the terminologies used and their precise meaning between RWD sources and FDA-supported data standards…"	Because there might also be regional variations in data definitions across RWD sources, Janssen requests that FDA provide discussion of the impact that regional variations may have on RWD mapping.
Lines 147-155	In describing the Study Data Reviewer's Guide, which serves as the rationale for choosing CDISC elements, guidance indicates that sponsors should provide "a description of the general approach and anticipated impact of data mapping" as part of the document.	Janssen requests FDA to include recommendations for possible RWD mapping strategies. Janssen requests that that FDA include additional information such as what information should be included whether the Study Data Reviewer's Guide should be focused on data mappings with greater impact by defining what impact means and on what domains an 'impact' should be evaluated. Additionally, including a list of key elements, or headings, for the Study Data Reviewer's Guide would strengthen the guidance.
Line 151-153	In this section FDA references a "data dictionary that documents the definition of every data element used and all relevant information about the element".	Janssen requests that FDA clarifies the criticality of a data element to the conclusions of the study and to provide more information for critical elements (e.g., elements related to the outcome and exposure variables for the study) than for non-critical elements.
Lines 159-166	In this section FDA discusses considerations for data transformation and the examples of FDA-supported data standards. FDA also outlines challenges when transforming RWD.	The present draft guidance does not reference how to handle discrepant data from RWD source. Given that discrepant data cannot be queried the same way as we typically do in clinical trials, Janssen suggests documenting the rules that are applied for decision making as to what data to use in analysis and the openness for different expectations for structured versus unstructured data fields. Furthermore, additional challenges may arise for submission of probabilistic data elements generated by machine learning (e.g., a probability that a patient experienced some condition).

Glossary		
Lines 216, 230	Throughout the guidance, the FDA references CDISC's Glossary.	Janssen requests the following edits for harmonization of the current draft guidance's glossary with CDISC's Glossary: - for Row 216, "**Mapping**: in the context of representing or exchanging data, connecting an item or symbol to a code or concept. Compare to translation." - for Row 230 "**Non-interventional (observational) study**: A study where the medicinal product(s) is (are) prescribed in the usual manner in accordance with the terms of the marketing authorization. The assignment of the patient to a particular therapeutic strategy is not decided in advance by a trial protocol but falls within current practice and the prescription of the medicine is clearly separated from the decision to include the patient in the study. No additional diagnostic or monitoring procedures shall be applied to the patients and epidemiological methods shall be used for the analysis of collected data." Janssen also notes that it would be helpful if the definition of "noninterventional (observational) study also be considered in the context beyond single "drug of interest" (e.g., to include observational studies of exposures/treatments, etc.).
Line 185	In this section FDA defines data curation.	Janssen suggests changing "data curation" to "**data capture**" and to refer CDSIC Glossary for definition.
Line 190	In this section FDA defines data domain, which is very specific to study examples, whereas is also used in RWD.	For clarity, Janssen suggests creating a more generic definition to cover both study and RWD examples. The following language is suggested: "Data Domain: a collection of logically related observations (with a common, specific topic) that are normally collected for all subjects in a clinical investigation **or as part of routine care**. NOTE: The logic of the relationship may pertain to the scientific subject matter of the data or to its role in the trial/study. Example domains include laboratory test results, adverse events, concomitant medications

Line 196	In this section FDA defines data standards.	Janssen recommends the following edit: "**Implementation of** data standards make submissions predictable, consistent, and have a form that an information technology system or a scientific tool can use." and to provide clarification to which scientific tools it refers.