Submitting Study Datasets for Vaccines to the Office of Vaccines Research and Review

---

Guidance for Industry

Technical Specifications Document

This guidance is for immediate implementation.

FDA is issuing this guidance for immediate implementation in accordance with 21 CFR 10.115(g)(4)(i). Submit one set of either electronic or written comments on this guidance at any time. Submit electronic comments to https://www.regulations.gov. Submit written comments to the Dockets Management Staff (HFA-305), Food and Drug Administration, 5630 Fishers Lane, Rm. 1061, Rockville, MD 20852. You should identify all comments with docket number FDA-2018-D-1216.

Additional copies of this guidance are available from the Office of Communication, Outreach and Development (OCOD), 10903 New Hampshire Ave., Bldg. 71, Rm. 3128, Silver Spring, MD 20993-0002, or by calling 1-800-835-4709 or 240-402-8010, or email ocod@fda.hhs.gov, or from the Internet at https://www.fda.gov/vaccines-blood-biologics/guidance-compliance-regulatory-information-biologics/biologics-guidances.

For questions on the content of this guidance, contact OCOD at the phone numbers or email address listed above.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Biologics Evaluation and Research
October 2019

U.S. FOOD & DRUG ADMINISTRATION

Submitting Study Datasets for Vaccines to the Office of Vaccines Research and Review

---

Guidance for Industry

Technical Specifications Document

Revision History

Date	Version	Summary of Revisions
April 2018	1.0	Initial Version
October 2019	2.0	Revised Language in Section 3.5 and 4.0 Added Language in Section 2.0, 3.2, 3.3, 4.0 and 7.9 Added Section 6.0

Table of Contents

Contains Nonbinding Recommendations

Submitting Study Datasets for Vaccines to the Office of Vaccines Research and Review

Guidance for Industry

Technical Specifications Document

This guidance represents the current thinking of the Food and Drug Administration (FDA or Agency) on this topic. It does not establish any rights for any person and is not binding on FDA or the public. You can use an alternative approach if it satisfies the requirements of the applicable statutes and regulations. To discuss an alternative approach, contact the FDA staff responsible for this guidance as listed on the title page.

1.0 INTRODUCTION

This technical specifications document provides you, sponsor/applicant detailed information and specifications for the content of datasets submitted to FDA's Center for Biologics Evaluation and Research (CBER) Office of Vaccines Research and Review (OVRR) and is designed to aid clinical and statistical reviewers in the review of vaccine applications, e.g., biologics license applications. We, FDA, recommend submission of these datasets as part of the applicant's data tabulation datasets. The submission of standardized datasets facilitates review and analyses of the data and allows for pooling of data, when appropriate. These specifications reflect current CBER OVRR thinking, are built to be consistent with the FDA Study Data Technical Conformance Guide[1], and are generally consistent with the Therapeutic Area User Guide (TAUG) for Vaccines[2].

If there is a question regarding a specific submission or implementation of a particular data standard, the sponsor should contact the appropriate review division in OVRR for specific submission questions or the appropriate contact for data standards issues at cber-edata@fda.hhs.gov. For more general recommendations on the use and submission of standardized trial data, the sponsor should refer to the FDA Study Data Technical Conformance Guide.

FDA's guidance documents, including this guidance, do not establish legally enforceable responsibilities. Instead, guidance describes the FDA's current thinking on a topic and should be viewed only as recommendations unless specific regulatory or statutory requirements are cited. The use of the word *should* in FDA's guidance means that something is suggested or recommended but not required.

[1] https://www.fda.gov/media/122913/download
[2] https://www.cdisc.org/standards/therapeutic-areas/vaccines

2.0 SUBMISSION OF DATA FOR VACCINE TRIALS

The domains to be utilized in all Study Data Tabulation Model (SDTM) datasets submitted to OVRR include: Trial Summary (TS), Demographics (DM), Subject Visits (SV), Concomitant Medications (CM), Exposure (EX), Disposition (DS), Protocol Deviations (DV), and Medical History (MH). If a physical exam is included in the protocol, the information should be included in the Physical Exam (PE) domain. Similarly, if safety labs are included in the protocol, the information should be included in the Laboratory Test Results (LB) domain. Usage of the Adverse Event (AE), Clinical Event (CE), Findings About (FA), Vital Signs (VS), and Microbiology Specimen (MB) domains will be detailed below.

The DS domain should always include the Standardized Disposition Name (DECOD), Reported Term (TERM), Category (CAT), and Start Date/Time of Event (STDTC) variables. If relevant (i.e., multiple doses administered over time), the DS domain should include the use of the Order of Element within Arm (TAETORD) timing variable. The TAETORD timing variable should also be included in other domains when the vaccine is multi-dose.

In addition, the domains and variables to be utilized for safety, immunogenicity, and efficacy (i.e., clinical disease endpoint) data are detailed below.

3.0 SUBMISSION OF SAFETY DATA FOR VACCINE TRIALS

Safety data submitted for vaccine clinical trials should include reactogenicity data (i.e., a set of prespecified AEs collected within a prespecified time frame, often referred to as solicited AEs or reactions), unsolicited AEs, medically attended adverse events (MAAEs), and death. The sections below will detail the usage of domains and variables expected in submissions to OVRR. Any usage of domains and variables that differ from these recommendations will need to be agreed upon prior to submission of datasets.

3.1 Reactogenicity Data

Reactogenicity data should be represented primarily in the CE domain with the Findings About Clinical Event (FACE) domain and VS domain to provide the specific information for each event. A more detailed description of how reactogenicity data should be collected is described in the Vaccine TAUG2. We prefer that the Vaccine TAUG "flat model" be utilized, i.e., that data for each day be included, even if a subject never experienced a particular event. However, the "nested model," which includes only a summary record for a particular event if a subject never experienced that event, may be necessary for large trials with significant amounts of data. Sponsors should discuss which model is appropriate with their review team prior to beginning their clinical trial.

In the flat model, the CE dataset should include a record for each event for each subject, with the Clinical Event Reported Term (CETERM) as the event name and the Clinical Event Occurrence (CEOCCUR) = "Y" or "N." CEOCCUR is "null" if the reactogenicity event was not measured on a particular day (i.e., data are missing) and none of the other assessment days for that subject are "Y." If an event occurred, the clinical event start day/date (CESTDY/CESTDTC) and end day/date (CEENDY/CEENDTC) of the reactogenicity event should be collected and included in the dataset. If it is unknown

when an event started and/or ended (e.g., because the diary card was not completed), the variable(s) CESTDTC and/or CEENDTC should be null, and a comment should be provided in the Comments (CO) domain. The pattern of the event (i.e., continuous or intermittent) can be noted by using the Clinical Event Pattern of Events (CEPATT) variable, if collected. In both cases, the daily records are represented in FACE. All daily temperature measurements are represented in the VS domain. Standard units should either be Celsius (°C) or Fahrenheit (°F), but not both within a dataset. A dataset-level relationship in Related Records (RELREC) should be used to represent relationships between CE and FACE and between VS and assessments of fever in CE.

The annotated Case Report Form (CRF) should include a check box to indicate whether the reactogenicity event did or did not occur, during the prespecified time frame. A check box in the annotated CRF should also capture whether the reactogenicity event was collected every day, during the assessment interval for a subject. If the Occurrence (OCCUR) variable is null, the Completion Status (--STAT) and Reason Not Done (--REASND) variables in FACE and VS and CE (covering the assessment interval) should be utilized.

Reactogenicity events should be indicated by using "reactogenicity" in the variable Clinical Event Category (CECAT). It should be further subcategorized into "administration site" or "systemic," using the variable Clinical Event Subcategory (CESCAT).

If a reactogenicity event should happen to continue beyond the assessment interval, it should also be represented in the AE domain, with the CE domain record indicating that the event extended beyond the assessment interval, using the variables Clinical Event End Reference Time Point (CEENTPT) (e.g., "Day 7") and Clinical Event End Relative to Reference Time Point (CEENRTPT) ("Ongoing"). A check box in the CRF, indicating whether the reactogenicity event is ongoing after the assessment period, is recommended. The event should be categorized in the Adverse Event Category (AECAT) variable as "reactogenicity." The start day/date (--STDY/--STDTC) and the end day/date (--ENDY/--ENDTC) of the reactogenicity event should be identical in both the CE and AE domain; whereas the duration (--DUR) should report the time that the event occurred as part of the assessment interval and as part of the continuance separately (e.g., an event that lasted 6 days in the assessment interval and 3 days beyond the assessment would be reported as Clinical Event Duration (CEDUR) = 6 days and Adverse Event Duration (AEDUR) = 3 days). We recommend one or more check boxes in the CRF, indicating the duration of the reactogenicity event in the assessment period and the duration of the reactogenicity event beyond the assessment period. If a reactogenicity event rises to the level of a Serious AE (SAE) during the prespecified assessment interval, it should be reported in the AE domain, and the Serious Adverse Event (AESER) variable should be utilized, as well as the CESER variable. The event should be categorized in the AECAT as "reactogenicity." The day and date that the reactogenicity event becomes serious should be reported in supplemental AE (suppae) domain Qualifier Variable Names (QNAM) = "SAEDY" and "SAEDTC," Qualifier Variable Labels (QLABEL) = "SAE day" and "SAE date," and Supplemental Qualifier Values (QVAL) = day/date that reactogenicity event became serious. A dataset-level relationship in RELREC should be used for each case. We prefer that, in both cases, FACE (and/or VS) be utilized to provide the day-to-day data for the reactogenicity event from start to end.

If a reactogenicity event is assessed by the investigator as unrelated to the vaccine, but instead related to another diagnosis (e.g., Viral Infection), the definitive diagnosis would be recorded in the AE domain with the day-to-day signs and symptoms reported in the Findings About Adverse Event (FAAE) domain. However, the reactogenicity event(s)) (i.e., symptoms), would still be reported in the CE (and FACE/VS) if obtained from the diary because the event(s) are prespecified and occur during the solicited assessment period. In this case CEREL (causality) can be reported as "N" as there is a plausible reason for the event that is not due to vaccine administration. AEREL would also be reported as "N". The relationship between AE and CE in the subject should be established using RELREC.

Please note it would be considered recall data, if you are reporting selected AE preferred terms (PT) from the CRF that are synonymous with solicited local administration/systemic reactogenicity events which "occurred" during the pre-specified reporting time period post-vaccination but were actually collected from the investigator in a follow-up visit. If you intend to report data based on recall (and not obtained from the diary), please minimize potential bias by including a reconstructed data flag in supplemental CE (suppce) using QNAM = "RECON," (QLABEL) = "reconstructed data" and (QVAL) "Y" in your dataset. Reconstruction should be limited to investigators who are still blinded to treatment group in a randomized, controlled trial using a scripted query for parameters (but perhaps not grading severity) if applied during a specified time period post-vaccination. A "reconstructed data" flag might support a comparison of results generated based on different types of data collection (e.g. diary card vs eCRF) as well as overall compiled safety data; and may be used to determine if randomization was preserved in "reconstructed" data to get a gross sense of whether the investigators were truly blinded as they queried subjects and whether all treatment groups had the expected representation in both the "purest" collection of the safety data (i.e., directly from the diary) and the "reconstructed" as well as whether there were differences between sites. Any plans to include reconstructed data in your submission should be discussed with the review team prior to submission.

If an analgesic/antipyretic is used for prevention or treatment of pain and/or fever associated with vaccine administration, please report this in the CM domain using the appropriate variables to report the date and time in ISO 8601 format or other character format with reference time point. The CMCAT (Concomitant Medication Category) should be used to differentiate "Prevention" versus "Treatment." Additionally, you should flag if analgesic/antipyretic administered (Y/N) using the variable CECONTRT (Clinical Event Concomitant or Additional Treatment Given).

A telephone contact that is used to ensure diary reporting compliance and/or to ascertain if any unsolicited AEs have occurred is considered a subject visit. Please report the study day of start of visit (SVSTDY), the visit number (VISITNUM) and the visit name (VISIT) as "telephone contact 1 (or 2 or 3, etc.)" for each subject.

3.2 Unsolicited Adverse Events (AEs)

Unsolicited AEs should be represented in the AE domain, regardless of whether they occurred during the prespecified assessment interval or after the prespecified assessment interval. The clinical details of the AE should be represented in the FAAE domain and possibly in the VS, CM, Procedures (PR), Healthcare Encounter (HO), and Death Detail (DD) domains. The following variables should be utilized in the AE domain: --SEV (Severity) (or –TOXGR (Standard Toxicity Grade)), --LLT (Lowest Level Term), --LLTCD (Lowest Level Term Code), --HLT (High Level Term), --HLTCD (High Level Term Code), --HLGT (High Level Group Term), --HLGTCD (High Level Group Term Code), --SOC (Primary System Organ Class) and --SOCCD (Primary System Organ Class Code). Additionally, you should report in the AETRTREF (Time Point Reference) variable the reference time point of "Dose 1, 2, or 3 etc." and AEELTM (Planned Elapsed Time from Time Point Reference) variable to indicate the beginning of the event (in days, hours/minutes if necessary) from the last vaccination administered (referenced as dose 1, 2, or 3 etc.).

In the AE domain, there should be one record per AE per subject for each unique event. The AE record reported by the investigator is submitted as a summary record, "collapsed" to the highest level of severity, causality, seriousness, and final outcome. The variables for study day/date should be utilized (AESTDY/AESTDTC and AEENDY/AEENDTC). The FAAE domain should be used to report the day-to-day information, concerning the AE. The CM, PR, HO, and/or DD domains should be utilized if the corresponding variables in AE are marked as "yes." A dataset-level relationship in RELREC should be provided.

3.3 Laboratory Safety Assessments

Laboratory safety assessment of clinical chemistry, hematology, and urine should be included in the LB domain. If the laboratory result for a particular assessment for a subject is outside the normal range, it should additionally be reported in the AE domain, with the highest level noted. A dataset-level relationship in RELREC should be used.

3.4 Medically Attended Adverse Events (MAAEs)

MAAEs should be reported in the AE and HO domain, and if needed, additional data should be reported in the FAAE, VS, and/or DD domains.

In certain circumstances, OVRR may request that a clinical trial protocol include collection of data relating to specific MAAEs, particularly autoimmune or auto-inflammatory diseases, among subjects in all treatment groups through 12 months or longer, following the last study vaccination. Since the request for this information is based on a theoretical potential, CBER considers the occurrence of a vaccine-related potential immune-mediated medical conditions (PIMMC) to be unexpected.

When a particular MAAE is noted as a new onset of a chronic disease (NOCD) or a PIMMC it should be reported in the AE domain and categorized in CAT with "NOCD" or "PIMMC," respectively.

If a chronic disease is exacerbated, following a vaccination, the MH domain should contain information on the disease, prior to the first day of vaccination, and the AE domain should contain information on the disease, following vaccination. The CAT and Subcategory (-SCAT) variables should be populated in both domains with a categorization/subcategorization of the chronic disease, e.g., kidney failure would be categorized as "renal medical history" and subcategorized as "renal failure." A dataset-level relationship in RELREC should be used.

3.5 Deaths

When a death occurs, it should be reported in the AE domain with supplemental death detail information provided in the DD domain. The DM domain should also indicate that a death has occurred, using the flag variable (DTHFL) and date/time of death variable (DTHDTC).

4.0 SUBMISSION OF CLINICAL DISEASE ENDPOINT EFFICACY DATA FOR VACCINE TRIALS

For clinical trials in which efficacy against clinical disease is an objective, efficacy data will primarily be reported in the CE domain with specific information provided in the MB domain, FACE domain, VS domain, and PE domain (if necessary).

To differentiate an efficacy clinical event (i.e., case of the disease of interest) from other clinical events, the variable CECAT should be utilized, and "efficacy" should be stated. Since efficacy is pre-specified, the variables CEOCCUR, Clinical Event Completion Status (CESTAT), and Clinical Event Reason Not Collected (CEREASND) should be included. The CETERM for the efficacy clinical event should be indicative of the disease, e.g., for an influenza vaccine, "influenza-like illness or ILI;" or for a human papillomavirus vaccine, "cervical intraepithelial neoplasia or CIN. Test results from assays conducted to confirm the presence of the microbe of interest (e.g., polymerase chain reaction (PCR), enzyme-linked immunosorbent assay (ELISA), cell culture, etc.) should be reported in the MB domain. If the disease of interest is confirmed (e.g., by the clinical investigator or by an adjudication committee, using the information from a clinical assessment and the results of confirmatory tests), it should be reported in CE variable CEDECOD (e.g., Influenza), and flagged in a supplemental DM domain with QNAM = "CDECASE," (QLABEL) = "clinical disease endpoint case flag," and (QVAL) = "Y."

If the disease of interest is not confirmed by either the clinical investigator or by an adjudication committee, the results should be flagged in a supplemental DM domain with QNAM = "CDECASE," QLABEL = "clinical endpoint case flag," and QVAL = "N." The manner in which the occurrence of a disease of interest is confirmed (e.g., by an adjudication committee or by the clinical investigator or by another mechanism) should be clearly delineated in the clinical protocol. A check box in the CRF, indicating the status of the disease of interest, is recommended.

If hospitalization occurs during the clinical trial because of occurrence of the "disease of interest," the CE domain variable Clinical Event Requires or Prolongs Hospitalization (CESHOSP) should be marked with "Y." Additional data regarding the hospitalization and/or other healthcare encounter should be provided in the HO domain.

If concomitant or additional treatment is prescribed to alleviate the severity of the disease of interest, the CE domain variable Clinical Event Concomitant or Additional Treatment Given (CECONTRT) should be utilized.

Variables available in AE that are also available for use in CE consist of Serious (SER), Outcome of Event (OUT), Involves Cancer (SCAN), Congenital Anomaly or Birth Defect (SCONG), Persistent or Significant Disability/Incapacity (SDISAB), Is Life Threatening (SLIFE), Results in Death (SDTH), and Occurred with Overdose (SOD).

5.0 SUBMISSION OF IMMUNOGENICITY DATA FOR VACCINE TRIALS

Data related to the assessment of immunogenicity should be reported in the Immunogenicity Specimen (IS) domain. The LB domain should not be used for immunogenicity data.

6.0 SUBMISSION OF COMCOMITANT MEDICATION DATA FOR VACCINE TRIALS

Previous immunizations should be reported in the CM domain. You should include the variable CMPRESP along with the accompanying variables -OCCUR, -STAT and -REASND, if the particular immunization is prespecified.

7.0 SUBMISSION OF DATA FOR MATERNAL IMMUNIZATION TRIALS

The relationship between the mother and child should be represented in Relationship of Study Subjects (RELSUB). To allow for the possibility that a mother is treated before the number of fetuses is known, fetuses are given a Unique Subject Identifier (USUBJID) value that appends an A or B to the mother's USUBJID.

In the DM domain, the informed consent and vaccination administration date for the fetus are the same as for the mother. The age of enrollment will be a negative number in weeks. The birth date for the infant should be reported, using the variable Date/Time of Birth (BRTHDTC).

Since the amount of vaccine that reaches the fetus is unknown, the Dose Description (DOSTXT) variable in EX is utilized, and "Fetal Exposure" should be noted instead of populating the variable Dose (DOSE).

The reproductive findings (RP) domain should be used to report the maternal date of last menstrual period, and from this, the gestational age at exposure can be calculated and reported in RP.