Assumptions in this section are applicable to Interventions, Events, and Findings class domains and will be used with domain-specific assumptions as appropriate.

The following assumptions will be implemented for **Interventions** class domains.

<table>
<thead>
<tr>
<th>Num</th>
<th>Collection Variable Use</th>
<th>Implementation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>--YN</td>
<td><img src="image" alt="Variables with the question text &quot;Were there any interventions?&quot; (e.g., &quot;Were there any concomitant medications?&quot;)" /></td>
</tr>
</tbody>
</table>

- Variables with the question text "Were there any interventions?" (e.g., "Were there any concomitant medications?") support the cleaning of data and confirmation that there are no missing values.

- Values collected for these fields will not be represented in subsequent tabulation datasets.
Categories and subcategories are determined per protocol design and values are generally not entered via CRF.

- Implementers may:
  - Prepopulate and display category values to help individuals involved in data collection understand what data should be recorded on the CRF.
  - Prepopulate hidden variables with the values assigned within their operational database.
  - Populate values directly in the tabulation dataset during dataset creation.
Variables for date and time

- The time an intervention started will be collected if there is a scientific or regulatory reason to collect this level of detail and the time can be realistically determined.
- Collection variables for date (e.g., --DAT, --STDAT, --ENDAT) will be concatenated with collection variables for time (e.g., --TO, --FROM, --JIRA) as applicable to populate tabulation variables for dates (e.g., --DTC, --STDTC, --ENDTC) using ISO 8601 format.
• --REASND is used with tabulation variable --STAT.

• The value "NOT DONE" in --STAT indicates that the subject was not questioned about the intervention or that data were not collected; it does not mean that the subject had no interventions.

• --SPID may be populated by the applicant's data collection system.
• If collected, --SPID it can be used as an identifier in a data query to communicate clearly to individuals involved in data collection the record in question.
- When free-text intervention treatments are recorded, the location may be included in the --TRT variable to facilitate coding (e.g., lung biopsy). Location may be collected when the applicant needs to identify the specific anatomical location of the intervention. This location information does not need to be removed from the verbatim --TRT when creating tabulation datasets.
- The non-standard variables --ATC1 through --ATC5 and --ATC1CD through --ATC5CD are used only when the intervention is coded using the World Health Organization’s Anatomical Therapeutic Chemical (ATC) classification system ([https://www.who.int/medicines/regulation/](https://www.who.int/medicines/regulation)): 1 = the anatomical main group, 2 = the therapeutic main group, 3 = the therapeutic/pharmacological subgroup, 4 = chemical/therapeutic/pharmacological subgroup, 5 = chemical substance.
- Implementers may also add MedDRA coding elements as NSVs to the Interventions domain if that dictionary is used for coding.
Applicants may collect location data using a subset list of controlled terminology on the CRF.

Applicants may prepopulate hidden variables with values assigned within their operational database.

There is currently some overlap across controlled terminology for --LOC, --LAT, and --DIR. While the overlap exists, ensure that this overlap is not part of database design.

The following assumptions will be implemented for Events class domains.

<table>
<thead>
<tr>
<th>Num</th>
<th>Field or Variable</th>
<th>Guidance</th>
</tr>
</thead>
</table>
| 1   | --YN              | • Variables with the question text "Were there any <events>?" (e.g., "Were there any adverse events?") support the cleaning of data and confirmation that there are no missing values.  
• These questions can be used on any CRF.  
• Values collected for these fields will not be represented in subsequent tabulation datasets. |
2  **--CAT, --SCAT**
- Categories and subcategories are determined per protocol design and values are generally not entered via CRF.
- Implementers may:
  - Prepopulate and display category values to help individuals involved in data collection understand what data should be recorded on the CRF.
  - Prepopulate hidden variables with the values assigned within their operational database.
  - Populate values directly in the tabulation dataset during dataset creation.

3  **Variables for date and time**
- The time of an event will be collected if there is a scientific or regulatory reason to collect this level of detail and the time can be realistically determined.
- Collection variables for date (e.g., --DAT, --STDAT, --ENDAT) will be concatenated with collection variables for time (e.g., --TIM, --STTIM, --ENTIM) as applicable to populate tabulation variables for dates (e.g., --DTC, --STDTC, --ENDTC) using ISO 8601 format.

4  **--OCCUR**
- --OCCUR may be used when a specific event is solicited (preprinted) on the CRF and the CRF uses an applicant-defined codelist.
- --OCCUR may be implemented while also allowing for a "NOT DONE" response.

5  **--REASND**
- --REASND is used with tabulation variable --STAT.
- The value "NOT DONE" in --STAT indicates that the subject was not questioned about the event or that data were not collected; it does not mean that the subject had no events.
### 6 SPID

- SPID may be populated by the applicant's data collection system.
- If collected, SPID it can be used as an identifier in a data query to communicate clearly to individuals involved in data collection the record in question.

### 7 Coding

- The collection variables used for coding are not data collection fields that will appear on the CRF. Applicants will populate values through the coding process.
- When free-text event terms are entered, the location may be included in --TERM to facilitate coding and further clarify the event. This location information does not need to be removed from the verbatim term when creating tabulation datasets.
- The CDASH variables --LLT, --LLTCD, --PTCD, --HLT, --HLTCD, --HLGT, --HLGTCD, --SOC, and --SOCCD are only applicable to events coded in MedDRA.

### 8 Location (--LOC, --LAT, --DIR, --PORTOT)

- Location is collected when the applicant needs to identify the specific anatomical location of the event.
- Applicants may collect location data using a subset list of controlled terminology on the CRF.
- Applicants may prepopulate hidden variables with values assigned within their operational database.
- There is currently some overlap across controlled terminology for LOC, LAT, and DIR. While the overlap exists, ensure that this overlap is not part of database design.

The following assumptions will be implemented for Findings class domains.
<table>
<thead>
<tr>
<th>Num</th>
<th>Field or Variable</th>
<th>Guidance</th>
</tr>
</thead>
</table>
| 1   | --CAT, --SCAT    | • Categories and subcategories are determined per protocol design and values are generally not entered via CRF.  
• Implementers may:  
  ◦ Prepopulate and display category values to help individuals involved in data collection understand what data should be recorded on the CRF.  
  ◦ Prepopulate hidden variables with the values assigned within their operational database.  
  ◦ Populate values directly in the tabulation dataset during dataset creation. |
| 2   | --PERF, --STAT, --REASND | • --PERF defines - variables to record whether an assessment has been performed/colllected. --REASND is used to collect a reason why an assessment was not done.  
• --PERF has the Question Text "[Were any/Was the] [--TEST/ topic] [measurement(s)/test(s) /examinations (s)/specimen(s) /sample(s) ] [performed/collected]?” are intended to assist in the cleaning of data and in confirming that there are no missing values.  
• --PERF may be used at the page, panel, or question level.  
• --PERF may be used during the creation of tabulation datasets to derive a value into the SDTM variable --STAT. The implementer can use a combination of --CAT, --SCAT, with the --TESTCD= "--ALL" and --TEST= "<Name of the CRF module>" to represent what tests were not performed.  
• Applicants must decide how to model each test not performed (e.g., to denote that all tests were not performed using TESTCD = "--ALL").  
• --STAT has the Question Text "Was the [--TEST] not [completed/answered/done/assessed/evaluated]?; Indicate if (the [--TEST] was) not [answered/assessed/done/evaluated/performed]." This is intended to be used to collect a simple "NOT DONE" check box at the page, panel, or question level. |
--REASND is used with SDTM variable --STAT only. The value NOT DONE in --STAT indicates that the findings test was not performed.

3  --SPID

- --SPID may be populated by the applicant's data collection system. If collected, it can be beneficial to use an identifier in a data query to communicate clearly to the site the specific record in question.
- This field may be populated by the applicant's data collection system.

4  Variables for date and time

- Time will be collected if there is a scientific or regulatory reason to collect this level of detail and the time can be realistically determined.
  - Metadata tables generally include --DAT and --TIM will be added from the CDASH Model as appropriate.
- Collection variables for date and time (e.g., --DAT, --TIM) will be used to collect the date or date and time that the test was performed, or the specimen was collected. The start and end dates and times (e.g., for specimen collection) will be collected as appropriate.
- The date of collection of a test can be derived from the date of visit. In such cases, a separate date of observation field is not required to be present on the CRF.
- Date and time variables will not be used to collect dates that are the result of a tests. Test results will be collected using --ORRES.
In metadata tables, many of the Findings class domains are presented in a normalized structure (1 record for each test) similar to a tabulation dataset, even though many data management systems hold the data in a denormalized structure (1 variable for each test).

When implementing collection standards in a denormalized structure, create variable names for the Findings --TEST and/or --TESTCD values. To do this:

- Define the denormalized variable names using available CDISC Controlled Terminology for --TESTCD; or
- When a system allows more than 8-character variable names, the following naming convention can be used: <--TESTCD>_<--tabulation variable name> where --TESTCD is the appropriate CT for the test code (e.g., DIABP_VSORRES, DIABP_VSLOC).

In the horizontal (denormalized) setting, collection variables such as --PERF, --LOC, and --STAT can be collected once for the whole horizontal record and applied to all of the observations on that record, or collected per test using collection variables, such as <--TESTCD>_<--PERF. When tabulation datasets are created, any variables collected for the entire horizontal record will be mapped to each vertical record per tabulation guidance.

In the horizontal (denormalized) setting, an identifier (e.g., --GRPID) can be used to identify all --TESTCD for the same collection record. This supports mapping of data collected in a horizontal setting to tabulation datasets and creation of RELRECs.
Tests and original results

- The value in --TEST will be 40 characters or less.
- The corresponding codelist value for the short test name, 8 characters or less, will be populated in the tabulation variable --TESTCD.
- Variable --TESTCD should be used to create a variable name and --TEST be used as the Prompt on the CRF.
- Both --TESTCD and --TEST are recommended for use in the operational database.
- Variable --ORRES is used to collect test results or findings in the original units per controlled terminology.

- If results are modified for coding, the --MODIFY variable contains the modified text.
- Variables --ORNRLO and --ORNRHI and --NRIND are used when normal or reference ranges are collected for results.
- Standardization of the original results and/or normal/reference ranges will be performed during the creation of tabulation datasets.
Location variables are used to collect the location of the test.
- Applicants may collect location data using a subset list of controlled terminology on the CRF.
- Applicants may prepopulate hidden variables with values assigned within their operational database.
- There is currently some overlap across controlled terminology for LOC, LAT, and DIR. While the overlap exists, ensure that this overlap is not part of database design.
-ORRES, --RES, --DESC, and --RESOTH are used to collect results. It is recommended that:

- --ORRES is used when the result is collected using a single question. The result will map directly to the tabulation variable --ORRES.
- --RES and --DESC are used when a pair of questions are asked to collect the result; for example, the question "Is the <condition> [absent/present]?" with a follow-up question "What is the finding that was observed?"
- --RES and --RESOTH are used when a question is asked that allows the selection of a prespecified finding, with a follow-up question to ask about the pre-specified response "OTHER". For example, the question "What is the result?" with a set of prespecified responses, including the choice "OTHER" with the follow-up question "Specify, Other".

## Root variables

The Findings About Events and Intervention domains use the same root variables as the Findings domain, with the addition of the --OBJ variable.