



Study Data Tabulation Model: Conformance Rules User Guide

Version 1.0

**Prepared by the
SDS Sub-Team for SDTM Conformance Rules**



Revision History

Date	Version	Summary of Changes
2016-12-12	1.0 Final	

See [Appendix B](#) for Representations and Warranties, Limitations of Liability, and Disclaimers

CONTENTS

1	INTRODUCTION	3
2	PURPOSE.....	3
3	SPECIFICATION FOR THE RULES METADATA MODEL.....	3
4	ASSUMPTIONS FOR THE RULES METADATA MODEL.....	5
5	EXAMPLE RULES.....	6
	APPENDICES	8
	APPENDIX A: SDS SDTM CONFORMANCE RULES SUB-TEAM MEMBERS	8
	APPENDIX B: REPRESENTATIONS AND WARRANTIES, LIMITATIONS OF LIABILITY, AND DISCLAIMERS.....	9

1 Introduction

The CDISC Study Data Tabulation Model (SDTM) guides the organization, structure, and format of standard clinical trial tabulation datasets submitted to a regulatory authority.

This document recommends standard practices for use in formulating and documenting compliance and conformance rules as a supplement to the SDTM and SDTM Implementation Guide (SDTMIG). The guidance governs rules intended for application to human clinical trials only. Rules governed by this guidance are not assumed to be universally programmable, that is, capable of being implemented as automated checks.

2 Purpose

The purpose of this guide is to document a standard, concise structure for identifying and classifying SDTM and SDTMIG text that may constitute a conformance rule definition. The structure for the rules, the Rules Metadata Model, and the conventions for its content are described in detail. Users of this guide are aided to both understand rules in approved rules catalogs that implement this guidance and to author their own candidate rules (e.g. for submission with a Therapeutic Area User Guide) if the need arises.

3 Specification for the Rules Metadata Model

The following table provides a description of the columns in the SDTM Rules metadata:

Metadata	Description	Req
Rule ID	Identifier used to uniquely identify a rule. The ID is non-descriptive text string comprised of a simple prefix 'CG' with a 4-digit numeric suffix. The prefix is an abbreviation for 'Conformance Guidance' and is intended to denote the purpose of the rule as specific to governing compliance to guidance published by the CDISC organization, and as distinct from business and/or data validation rules. The numeric suffix is a unique sequential counter starting from '0001' that does not imply rule precedence or ranking based on a position in this sequential series.	Yes
Class	<p>Observation Class or rule category scope for the rule using an abbreviation of 3-characters or less:</p> <ul style="list-style-type: none"> • ALL – All observation classes • SPC – Special-Purpose Class • FND– Findings Class • EVT – Events Class • INT – Interventions Class • FNA – Findings About • TDM – Trial Design Domains • AP – Associated Persons Domains <p>When a rule applies to multiple classes or categories each is listed separated by commas. If the list is extensive then the use of the NOT() token (where the list of <i>excluded</i> Classes or categories is listed inside the parentheses [e.g., NOT(FND)]) is preferred. If ALL is used in the Class, then ALL should be the only value in this column.</p>	Yes
Domain	Domain scope for the rule using standard CDISC domain abbreviations, with exceptions being ALL or NOT(). When a rule applies to multiple domains, each domain is listed separated by commas. If the list of included domains is extensive, then the use of the NOT() token (where the list of <i>excluded</i> domains is listed inside the parentheses [e.g., NOT(SC,EG,PE)]) is preferred. If ALL is used in the Domain, then ALL should be the only value in this column.	Yes

Metadata	Description	Req
Variable	Variable scope for the rule. The variable name should be listed here except when the rule applies to multiple domains/observation classes. In such cases, a stubbed version of the variable name with dashes in place of the domain abbreviation should be used (e.g., --STDY, --ENDY). If the rule applies generally across many or all variables, then the keyword “GEN” should be the only value in this column.	Yes
Rule	A concise and unambiguous statement of the conformance or compliance principle to be applied. Rules represent and are stated as a standard to be met, as opposed to a failure condition or error statement. Only one principle is stated per defined rule. If the rule only applies during a circumstance or condition, then the condition is stated separately in the ‘Condition’ column.	Yes
Condition	Whenever a rule is applied only in a conditional set of circumstances the condition (or conditions) is defined in this column. Some guidelines for composing the condition statements are as follows: <ul style="list-style-type: none"> • Conditions are not prefixed with prepositions (e.g., ‘if’, ‘when’, ‘where’) as these are implicit.. • Multiple conditions are separated by standard logical operators (e.g., ‘and’, ‘or’, ‘and not’). • When referencing variables in the same domain, it is not necessary to use a domain prefix. However, references to variables in other domains must use the domain prefix (e.g., EG.VISIT, AE.AESER) • When referencing a character value, then enclosing quotes are required to make these distinct from variable references (e.g., “= ‘Y’” vs. “= ACTARMCD”). • A stubbed version of the variable name with dashes in place of the domain abbreviation are permitted and are required when the condition applies to multiple classes and/or domains (e.g., ‘--TRT’). • When referencing a variable that is part of a series of variables sharing the same root name but distinguished by a numeric index sequence e.g. TSVAL1, TSVAL2, TSVAL3, the following may apply: <ul style="list-style-type: none"> ○ A lowercase ‘n’ denotes the variable index number ○ Relative index notation e.g. ‘(n+1)’, ‘(n-1)’, indicates the position of one variable in the series relative to another. 	No
Guide	Indicates source document (e.g., SDTM Implementation Guide [IG vN.N] SDTM [Model vN.N]).	Yes
Section	If available, include specific section reference number/letter	No
Item	Within a domain model, values may be: Description/Overview, Specification, Assumption, or Example. Where applicable, (e.g., General Assumptions for All Domains), include specific item number as reference.	No
Cited Guidance	The text that states or implies the rule. It is recommended to include all directive statements for completeness (i.e., must, required, should, could, and may statements), and to copy/paste text from IG directly as needed. For Section 4, General Assumptions for All Domains, suggest representing once (vs. repeating per domain). Exceptions to be noted by domain.	No
IG Versions	SDTMIG versions for which the rule is in effect. Use standard numeric portion of SDTMIG version (e.g., 3.2, 3.3). If rule is in effect for multiple versions, each SDTM IG version number is listed, delimited by commas.	Yes
Batch ID	Identifier for the version or release of the rule. Used to track additions and updates to the rules catalog.	Yes
Programmable	Indicator that a rule may be able to be implemented as an automated check. Values are: <ul style="list-style-type: none"> • Y – Yes • C – Conditional. Rule may be programmable based on conditional factors. Some explanations of these factors are provided where applicable in the ‘Programmable Flag Comment’ column. 	Yes

Metadata	Description	Req
Programmable Flag Comment	Supplemental explanatory text for rules where there is a condition or factor as to whether they are able to be programmed as an automated check. In most cases this text would indicate a specific dependency on data or metadata that cannot be assumed to be always present and available.	No
FDA Rule ID	FDA Rules are referenced as an aid in understanding the context and impact of the SDTM rules. The FDA rules were in no way a source or an inspiration for the SDTM rules—these are based solely on guidance published by CDISC. The FDA Rule ID column header may be updated to indicate the referenced versions.	No

4 Assumptions for the Rules Metadata Model

1. General
 - a) The rule Class, Domain, Variable, Rule, and Condition elements collectively comprise the definition of the Rule. The Class, Domain, Variable, and Condition elements define the scope of the Rule's application.
 - b) All rules are stated uniquely at the highest applicable level of the rule hierarchy. For example, a rule stating that "ARM must not be null" (because its Core variable status is Required) should be implemented instead as a general rule for Required variables for all classes, all domains, and all variables, with a condition (e.g., Rule: " \wedge = null", Condition: "Core = REQ").
 - c) There is only one object or target variable for a rule. If the rule has an interdependent element (e.g., ARM must equal ACTARM when ARM = "Screen Failure"), then there should be two related rules where the object of the first is ARM and the second ACTARM.
 - d) Rules targeting permissible variables are only applicable if the variables are present in the dataset and populated with data.
 - e) A rule without a Condition applies universally to all variables within the scope defined by the Class, Domain, and Variable columns.
2. Rule Syntax
 - a) Observation classes, domains, and variable names are in uppercase.
 - b) Operators are in lowercase.
 - c) Any reference to a variable in a domain other than object variable's domain is in the form "Domain.Variable" (e.g., TA.ARM, DM.SEX).
 - d) Logical and mathematical operators are used in place of phrases such as 'less than or equal to', 'not greater than', 'should be', 'must equal', etc. This is to enforce a concise and unambiguous rendering of the rule. The following are examples of logical operator substitutions:

Use:	Instead of:
=	Should equal, must equal, equals
<=	Less than or equal to
<	Less than
>	Greater than
>=	Greater than or equal to
in	Must be defined in, must include
\wedge , not	Not – general negation operator

- e) If a variable has approved values described in guidance, but not in CDISC Controlled Terminology, then the syntax should be "Variable in ('value1', 'value2')". Quotes around individual text values in the list are required.

3. Rule Terminology

- a) Conventions in terminology are used to describe conditions or requirements in a standard way. The following table lists these terms and definitions:

Term	Definition
first	The first instance (e.g., date [implies sorting])
last	The last instance (e.g., date [implies sorting])
unique	Only occurrence of value within defined scope
present in dataset	Object is present in dataset, (does not imply that object is populated with data)
null	Object does not contain data
one-to-one	Value pair is a unique, isomorphic relation

- b) If a variable's approved values are a discrete list described in CDISC Controlled Terminology, then the syntax should be "Variable in {CT List Name}". Note the use of braces instead of parentheses.
- c) When describing a condition where a variable does not contain a value use the keyword "null" e.g., "AGETXT = null" rather than phrases like "is missing", "equals blank", "should not be populated".
- d) When describing a rule that has an expectation or a restriction to a certain type of data, the following standard terminology is used.

Data type	Definition
Numeric	Numbers, format chars (".", "+", or "-") only
Integer	Positive, whole numbers only
Alphabetic	Letters of alphabet only
Alphanumeric	Alphabetic and numeric format characters only
Boolean	True or False
Date	ISO 8601 standard datetime and date formats

4. Potential Rule Implementation as Automated Checks

- a) The 'Programmable' and 'Programmable Flag Comment' columns are the opinions of volunteer contributors and do not guarantee that the rules can be implemented as automated checks.
- b) Rules with 'Programmable' indicators of 'N' or 'C' are not to be considered less important than those marked with a 'Y'. Performance of manual checking by applying rules requiring human intervention and interpretation are an essential part of validating SDTM conformance.
- c) Rules are stated in the prescriptive, that is, they define what is correct. Therefore, it is important to keep in mind that it may be necessary to reinterpret the rules for systems that use failure criteria and that this may not be practical or advisable in some cases.

5 Example Rules

Example 1

Rules for use of --PRESF and --OCCUR

The example rules below illustrate a number of rule-metadata conventions, and general principles in defining the conformant use of the --PRESF and --OCCUR variables. Since these variables are permitted in only certain domains in the Events and Interventions observations classes, the rule metadata fields Class and Domain are used to define the scope accordingly. Notice the use of NOT() in the Domain column: this is used to define domains where this rule would not apply. This convention is useful when the list of applicable domains is cumbersome, and has the advantage of not precluding the rule from being extended to custom domains where it might apply. The Variable, Rule, and Condition fields used the '--' stub prefix to specify that the rules apply to the same root variables in multiple domains.

Rule CG0085 – Uses a delimited list of allowable values and the ‘null’ keyword in the rule. This rule does not have a Condition since it applies to every scenario where --PRESP can be used.

Rule CG0086 – Illustrates use of a compound condition with the ‘and’ operator.

Rule CG0088 – Shows the use of ‘not present’ that specifies that the --OCCUR variable should not be present in the dataset if the --PRESP variable is also not present. This ‘null’ keyword would not be used in this case as it would allow the presence of the variable with null values.

ID	Class	Domain	Variable	Rule	Condition
CG0085	EVT, INT	NOT(DS,DV,EX)	--PRESP	--PRESP in ('Y', null)	
CG0086	EVT, INT	NOT(AE,DS,DV,EX)	--OCCUR	--OCCUR ^= null	--PRESP = 'Y' and --STAT = null and --OCCUR is present in dataset
CG0087	EVT, INT	NOT(AE,DS,DV,EX)	--OCCUR	--OCCUR = null	'--PRESP ^= 'Y' and --OCCUR is present in dataset
CG0088	EVT, INT	NOT(AE,DS,DV,EX)	--OCCUR	--OCCUR not present in dataset	'--PRESP not present in dataset
CG0089	EVT, INT	NOT(AE,DS,DV,EX)	--PRESP	--PRESP = 'Y'	--OCCUR ^= null

Example 2

The following rules illustrate various types of conditions that would require the DM.DTHFL to be set to ‘Y’. Note the references to other domain objects in the condition statements.

ID	Class	Domain	Variable	Rule	Condition
CG0132	SPC	DM	DTHFL	DTHFL = 'Y'	SS.SSSTRESC = 'DEAD'
CG0133	SPC	DM	DTHFL	DTHFL = 'Y'	DD record present for subject
CG0134	SPC	DM	DTHFL	DTHFL = 'Y'	AE.AEOUT = 'FATAL'
CG0135	SPC	DM	DTHFL	DTHFL = 'Y'	AE.AESDTH = 'Y'
CG0136	SPC	DM	DTHFL	DTHFL = 'Y'	DS.DSDECOD = 'DEATH'

Appendices

Appendix A: SDS SDTM Conformance Rules Sub-Team Members

Team Member	Organization
Stetson Line	Independent
Carlo Radovsky	etera solutions
Tom Guinter	Pinnacle21
Anne Russotto	Independent
Sue Sullivan	D-Wise Technologies
Gary Walker	Quintiles
Daniel DiPrimeo	Puma Biotechnology
Naveed Khaja	Independent
Abhishek Dabral	Forest Laboratories
Manuel Anido	Forest Laboratories
Lynn Anderson	etera solutions
Max Kanevsky	Pinnacle 21
Helena Sviglin	FDA, CBER
Janet Reich	Amgen
Sergiy Sirichenko	Pinnacle 21
Nate Freimark	Theorem Clinical Research
Michael Beers	Pinnacle21
David Roulstone	Pinnacle21

Appendix B: Representations and Warranties, Limitations of Liability, and Disclaimers

CDISC Patent Disclaimers

It is possible that implementation of and compliance with this standard may require use of subject matter covered by patent rights. By publication of this standard, no position is taken with respect to the existence or validity of any claim or of any patent rights in connection therewith. CDISC, including the CDISC Board of Directors, shall not be responsible for identifying patent claims for which a license may be required in order to implement this standard or for conducting inquiries into the legal validity or scope of those patents or patent claims that are brought to its attention.

Representations and Warranties

“CDISC grants open public use of this User Guide (or Final Standards) under CDISC’s copyright.”

Each Participant in the development of this standard shall be deemed to represent, warrant, and covenant, at the time of a Contribution by such Participant (or by its Representative), that to the best of its knowledge and ability: (a) it holds or has the right to grant all relevant licenses to any of its Contributions in all jurisdictions or territories in which it holds relevant intellectual property rights; (b) there are no limits to the Participant’s ability to make the grants, acknowledgments, and agreements herein; and (c) the Contribution does not subject any Contribution, Draft Standard, Final Standard, or implementations thereof, in whole or in part, to licensing obligations with additional restrictions or requirements inconsistent with those set forth in this Policy, or that would require any such Contribution, Final Standard, or implementation, in whole or in part, to be either: (i) disclosed or distributed in source code form; (ii) licensed for the purpose of making derivative works (other than as set forth in Section 4.2 of the CDISC Intellectual Property Policy (“the Policy”)); or (iii) distributed at no charge, except as set forth in Sections 3, 5.1, and 4.2 of the Policy. If a Participant has knowledge that a Contribution made by any Participant or any other party may subject any Contribution, Draft Standard, Final Standard, or implementation, in whole or in part, to one or more of the licensing obligations listed in Section 9.3, such Participant shall give prompt notice of the same to the CDISC President who shall promptly notify all Participants.

No Other Warranties/Disclaimers. ALL PARTICIPANTS ACKNOWLEDGE THAT, EXCEPT AS PROVIDED UNDER SECTION 9.3 OF THE CDISC INTELLECTUAL PROPERTY POLICY, ALL DRAFT STANDARDS AND FINAL STANDARDS, AND ALL CONTRIBUTIONS TO FINAL STANDARDS AND DRAFT STANDARDS, ARE PROVIDED “AS IS” WITH NO WARRANTIES WHATSOEVER, WHETHER EXPRESS, IMPLIED, STATUTORY, OR OTHERWISE, AND THE PARTICIPANTS, REPRESENTATIVES, THE CDISC PRESIDENT, THE CDISC BOARD OF DIRECTORS, AND CDISC EXPRESSLY DISCLAIM ANY WARRANTY OF MERCHANTABILITY, NONINFRINGEMENT, FITNESS FOR ANY PARTICULAR OR INTENDED PURPOSE, OR ANY OTHER WARRANTY OTHERWISE ARISING OUT OF ANY PROPOSAL, FINAL STANDARDS OR DRAFT STANDARDS, OR CONTRIBUTION.

Limitation of Liability

IN NO EVENT WILL CDISC OR ANY OF ITS CONSTITUENT PARTS (INCLUDING, BUT NOT LIMITED TO, THE CDISC BOARD OF DIRECTORS, THE CDISC PRESIDENT, CDISC STAFF, AND CDISC MEMBERS) BE LIABLE TO ANY OTHER PERSON OR ENTITY FOR ANY LOSS OF PROFITS, LOSS OF USE, DIRECT, INDIRECT, INCIDENTAL, CONSEQUENTIAL, OR SPECIAL DAMAGES, WHETHER UNDER CONTRACT, TORT, WARRANTY, OR OTHERWISE, ARISING IN ANY WAY OUT OF THIS POLICY OR ANY RELATED AGREEMENT, WHETHER OR NOT SUCH PARTY HAD ADVANCE NOTICE OF THE POSSIBILITY OF SUCH DAMAGES.

Note: The CDISC Intellectual Property Policy can be found at http://www.cdisc.org/system/files/all/article/application/pdf/cdisc_20ip_20policy_final.pdf.