



PhUSE US Connect 2019

Paper DH06

7 Habits of Highly Effective (Validation Issue) Managers

Amy Garrett, Pinnacle 21, Blue Bell, USA

ABSTRACT

Pinnacle 21 Validator identifies problems in data; however, diagnostics, assessment and resolution of reported validation issues may feel like a complicated, never-ending process. In this presentation, we will discuss common challenges in managing data validation issues and how to handle them effectively. We will show you how to identify the source of validation issues, and how to classify them to understand when to fix or when to explain. We will also discuss cross-team collaboration, ways to improve your process, and habits that lead to faster issue resolution.

INTRODUCTION

All issues for a CDISC deliverable are not created equal: they come from different sources, have different impacts, and require different means of resolution. Programmers rely on Pinnacle 21 Validator to identify problems with data but resolving issues is often difficult - especially when there are several stakeholders to satisfy including your manager, the study statistician, the sponsor, and ultimately FDA and PMDA reviewers. An issue manager is usually a programmer, but it could be anyone at your organization; this paper talks about how you can be an effective issue manager that satisfies stakeholders by ensuring issues are resolved quickly and correctly.

Briefly, the seven habits are:

1. Validates early
2. Gathers all relevant info about issues
3. Identifies the source(s) of issues
4. Tracks changes between validation reports
5. Communicates issues to others
6. Knows when to fix validation issues
7. Knows when to explain validation issues

HABIT ONE: VALIDATES EARLY

Resolving validation issues is part of the CDISC process for ongoing studies. CDISC data are maintained many years in long clinical trials and there is a growing demand to have SDTM sooner than ever before, often leading programmers to have specifications and programs in place by the time a database goes live. When data conversion starts early, so should validation. Early validation can actually save time by acting as a first-pass QC to find and fix programming and specification-related issues. Using validation as a first-pass QC method will cut down on the amount of back-and-forth between primary and validation programmers later.

HABIT TWO: GATHERS ALL RELEVANT INFO ABOUT ISSUES

After you run a validation, it's important to be able to look-up erred records in more detail, sometimes even tracing data all the way back to the raw EDC data. One reason to do this is so that you can accurately identify the source of each issue. Another reason is because you need to communicate with data management about some issues and data managers often need more information about the record than what is provided on the Details tab of a Pinnacle 21 Validation Report. One way to quickly get more information about erred records is to use a SAS® macro like the one presented in *The Devil is in the Details – Reporting from Pinnacle 21 (OpenCDISC) Validation Report* (Garrett and Whalen, 2015). This macro creates a report of each issue on a separate excel tab and provides complete information about each erred record. Alternatively, Pinnacle 21 Enterprise provides full details about erred records directly within the system (see Figure 1).



Record	STUDYID	USUBJID	AESEQ	AETERM	AEDECOD	AESEV	AESER	AEACN	AESCONG	AESDISAB	AESDTH	AESHOSP	AESLIFE	AESMIE	AESTDTC
1576	NIDA-CTN-0001	01_067888	661	VOMITTING	VOMITTING NOS	SEVERE	N	NONE			Y				2001

Figure 1. Example of record details from Pinnacle 21 Enterprise

HABIT THREE: IDENTIFIES THE SOURCE(S) OF ISSUES

In order to understand where quality problems originate, it's important to identify the source of each issue and categorize it. Categories can also be used as a way to determine which issues need to be delegated to other people for resolution (for example, data collection issues should be assigned to data management). One way this is achieved is by assigning a primary source to each issue. In Pinnacle 21 Enterprise, users can make use of tags to categorize each issue by its source; the tag(s) assigned will be automatically exported to the validation report. Figure 2 is an example of tagging issues in Pinnacle 21 Enterprise. Alternatively, you can directly type the issue source into Excel validation report. Most issues will fall into one or more of the following categories:

- data collection error
- programming/spec error
- metadata (define.xml)
- sponsor-defined addition
- study is ongoing
- false positive

Issue ID	Description	Tags	Count	Delta	Severity	Action
SD0009	No qualifiers set to 'Y', when AE is Serious	Data Programming	157	+15	High	Open
SD0059	Define.xml/dataset variable type mismatch	Programming Define	2	0	High	Open
SD0080	AE start date is after the latest Disposition date	Data	15	0	High	Open
SD0091	AEOUT is not 'FATAL', when AESDTH='Y'	Data Spec Programming	60	+7	High	Open
SD1082	Variable length is too long for actual data	Programming	9	+3	High	Open
SD1076	Model permissible variable added into standard domain		1	0	Low	Open
SD1096	High risk of truncated value for AETERM variable	Programming	1	0	Low	Open
SD1201	Duplicate records in AE domain	Data Programming	7	-12	Low	Open
CM - Concomitant Medications (7 Issues)						
SD0035	Missing value for CMDOSU, when CMDOSE, CMDOSTXT or CMDOSTOT is provided	Data Programming	20	0	High	Open
CT2002	CMDOSU value not found in 'Unit' extensible codelist	Programming Define	1330	+34	Low	Open
SD0021	Missing End Time-Point value	Data Programming	6901	-78	Low	Open
SD0022	Missing Start Time-Point value	Data Programming	8	0	Low	Open
SD0031	Missing values for CMSTDT, CMSTRF and CMSTRPT, when CMENDTC, CMENRF or CMENRPT is provided	Data Spec Programming	5	0	Low	Open
SD0042	CMSTAT does not equal 'NOT DONE', when CMPRESP='Y' and CMOCCUR is NULL	Data Programming	341	0	Low	Open
SD1076	Model permissible variable added into standard domain		1	0	Low	Open

Figure 2. Example of Pinnacle 21 Enterprise issue table with tags applied to denote the issue source

HABIT FOUR: TRACKS CHANGES BETWEEN REPORTS

Validation is an ongoing process and it's important to track the delta from one validation report to the next so that you can determine when new issues are occurring and how quickly known issues are being resolved. There are three types of issues to track between validations: new issues, resolved issues, and issues that are on both reports (possibly affecting a different number of records). When the same issue is present in both reports, prior comments should be copied from one report to the next, as applicable. One way to track the delta between validation reports is to use a program (SAS macro or other application) to compare the two reports together and produce a consolidated report with comments carried forward. Another option is Pinnacle 21 Enterprise, which automatically copies issue comments from one validation report to the next and has a feature that allows users to easily compare any two versions of a validation report (see Figure 3).



Dataset	Rule	Message	Severity	Rate	Change	Risk	Status
GLOBAL	SD0061	Domain referenced in define.xml but dataset is missing	Warning	100%	-		Open
GLOBAL	SD1107	Missing LB dataset	Warning	100%	-		Open
GLOBAL	SD1108	Missing VS dataset	Warning	100%	-		Open
GLOBAL	TS0001	Missing LB dataset	Error	100%	-		Open
GLOBAL	TS0022	Missing VS dataset	Error	100%	-		Open

Figure 3. Example validation report comparison report filtered to only show new issues

HABIT FIVE: COMMUNICATES ISSUES TO OTHERS

The Pinnacle 21 Validation Report can serve as a communication device and be distributed to peers, managers and other stakeholders. Since all stakeholders may not be well-versed in CDISC, you should mark-up the validation report for better consumption. You should clearly state the source of an issue, who is responsible for resolving it, and provide any extra information that may be needed about an issue. Pinnacle 21 Enterprise allows users to assign issues to users, tag issues with their source, and provide comments about an issue, all of which are automatically exported to the Validation Report as depicted in Figure 4.

Pinnacle 21 Enterprise Validation Report									
Issue Summary									
Dataset	Rule ID	Publisher ID	Message	Severity	Found	Comments	Assigned	Tags	Explanation
AE	CT2001		AEACN value not found in 'Action Taken with Study Treatment' non-extensible codelist	Error	2362		Amy	mapping	
AE	CT2001		AEOUT value not found in 'Outcome of Event' non-extensible codelist	Error	2244		Amy	mapping	
AE	CT2001		AESEV value not found in 'Severity/Intensity Scale for Adverse Events' non-extensible codelist	Error	5		Amy	mapping	
AE	SD0008	FDAB003	Value for AEDECOD not found in MedDRA dictionary	Error	453	terms haven't been coded yet. (Amy, 1/9/2019 4:02 PM)	Cody	study design	
AE	SD0008C	FDAB017	Value for AEDECOD is in incorrect case	Error	1924	terms haven't been coded yet. (Amy, 1/9/2019 4:02 PM)	Cody	data collection	
AE	SD0012	FDAB034	AESTDY is after AEENDY	Error	5	One issue is related to SD0013 below, which cannot be fixed, but the other 4 are data issues (Amy, 1/9/2019 4:10 PM)	Brad	data collection	
AE	SD0013	FDAB034	AESTDTC is after AEENDTC	Error	1		Brad	data collection	This check fired for 1 records in AE where AESTDTC is after AEENDTC: USUBJID = XYZ-415, AESTDTC = 2011-09-05, AEENDTC = 2011-10. Though AEENDTC of '2011-10' is after AESTDTC of '2011-09-05', the dates cannot be compared because AEENDTC is a partial date. A complete date was not obtained prior to patient being lost-to-follow.
AE	SD0091		AEOUT is not 'FATAL', when AESDTH="Y"	Error	3	FATAL has been populated from the CRF - needs DM 3 to query site (Amy, 1/9/2019 4:10 PM)	Amy	data collection	
AE	SD1015		Invalid EPOCH	Error	147	EPOCH is null for these record - need to check programming logic (Amy, 1/9/2019 4:15 PM)	Amy	mapping	

Figure 4. Sample Pinnacle 21 Enterprise Validation Report with comments, assignee, tags, and explanations pre-populated based on information provided in the Enterprise system

HABIT SIX: KNOWS WHEN TO FIX VALIDATION ISSUES

An effective issue manager is one that knows when and how to fix each issue. You should understand the risk of each issue and prioritize fixing issues that have the highest impact on regulatory review. The FDA is no longer publishing severity, but it can still be used as a proxy for impact level; Errors and Reject rules usually have the highest impact on review. Pinnacle 21 Enterprise provides review impact for each issue, as show in Figure 5, and recommends that you fix issues with high review impact first. In general, data collection errors, programming/spec errors, and issues related to the define.xml (issue with DD prefix) should always be fixed while issues due to an ongoing study should resolve naturally overtime.

Pinnacle 21 Enterprise | Home | Report | Projects | Magic Pill | Studies | Sample Study | Download Report

Search

AE - Adverse Events (8 Issues)

SD0009: No qualifiers set to 'Y', when AE is Serious	Data	Programming	187	+15	High	Open
SD0059: Define.xml/dataset variable type mismatch	Programming	Define	2	0	High	Open
SD0080: AE start date is after the latest Disposition date	Data		13	0	High	Open
SD0091: AEOU is not 'FATAL', when AESDTH='Y'	Data	Spec	60	+7	High	Open
SD1082: Variable length is too long for actual data	Programming		9	+3	High	Open
SD1076: Model permissible variable added into standard domain			1	0	Low	Open
SD1096: High risk of truncated value for AETERM variable	Programming		1	0	Low	Open
SD1201: Duplicate records in AE domain	Data	Programming	7	-12	Low	Open

CM - Concomitant Medications (7 Issues)

SD0035: Missing value for CMDOSU, when CMDOSE, CMDOSTXT or CMDOSTOT is provided	Data	Programming	20	0	High	Open
CT2002: CMDOSU value not found in 'Unit' extensible codelist	Programming	Define	1330	+34	Low	Open
SD0021: Missing End Time-Point value	Data	Programming	8901	-78	Low	Open
SD0022: Missing Start Time-Point value	Data	Programming	8	0	Low	Open
SD0031: Missing values for CMSTDTC, CMSTRF and CMSTRPT, when CMENDTC, CMENRF or CMENRPT is provided	Data	Spec	5	0	Low	Open
SD0042: CMSTAT does not equal 'NOT DONE', when CMPRESP='Y' and CMOCUR is NULL	Data	Programming	341	0	Low	Open
SD1076: Model permissible variable added into standard domain			1	0	Low	Open

Figure 5. Sample issue summary table with Review Impact highlighted

Knowing how to fix each issue is challenging for any one individual, leading some companies to maintain a document with instruction guidance for how to fix the most common issues. This kind of document is especially helpful to more junior members of the team. An example of one such document can be found in the paper entitled *Common Pinnacle 21 Report Issues: Shall we Document or Fix* (Gupta, 2018). Pinnacle 21 Enterprise provides fix-tips for issues as well and can be customized for an organization (see Figure 6.)

SD0059: Define.xml/dataset variable type mismatch

Variable DataType in define.xml must match variable Type in dataset. Define-XML data types of 'integer' and 'float' match dataset type of 'Num', all other match 'Char'.

Common Source:	Programming, Define	P21 Fix Tips:
Review Impact:	High	
Dataset:	AE	
Affected:	2	
Assignee:	Michael Beers	
Status:	Open	<ul style="list-style-type: none"> Check the program and specs to verify type is set correctly. If so, then update the define.xml to match. Check the define.xml and verify type is correct. If so, update program/specification to match.

Figure 6. Same issue detail with Pinnacle 21 Fix tips provided

HABIT SEVEN: KNOWS WHEN TO EXPLAIN VALIDATION ISSUES



No study is perfect and there will always be a subset of pesky validation issues just won't go away. Any issue that cannot be fixed, even false positives, should be explained in the Reviewer's Guide. *Best Practice for Explaining Validation Results in the Study Data* describes a good explanation as "one that conveys transparency about the study data and increases the reviewability" (Kelly, 2018).

It's a best practice to explain issues consistently across an organization. To keep explanations consistent, some issue managers keep a list of standard explanations with bracket placeholders for study specific information and then copy these explanations into the Reviewer's Guide either manually or with a merge process. Another option is to use Pinnacle 21 Enterprise which allows users to upload standard explanations, apply them to issues and modify them as necessary. These explanations will automatically be exported to the Reviewer's Guide. Figure 7 shows an organization's standard explanation (right) and how it was applied to this study (left). Figure 8 shows how the explanation would appear in the Reviewer's Guide.

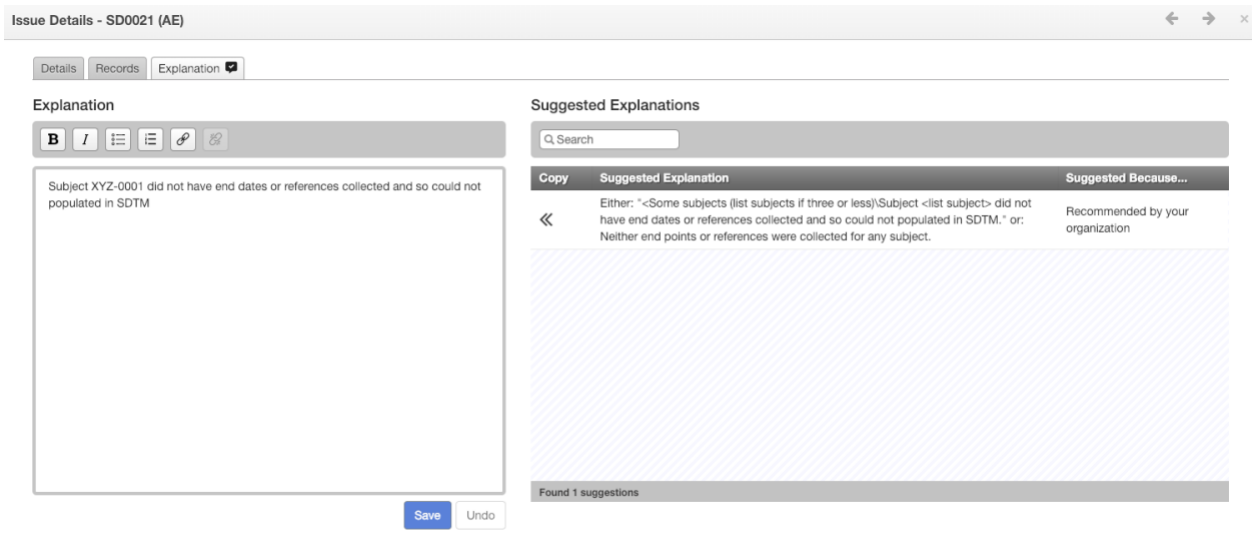


Figure 7. Example of Pinnacle 21 Enterprise Suggested Explanation and how a standard explanation can be customized with study-specific information.

Check ID	Diagnostic Message	FDA Severity	Dataset	Count (Issue Rate)	Explanation
SD0011	ARM is not 'Screen Failure', when ARMCD equals 'SCRNFAIL', or vice versa	Error	DM	25 (100.00%)	<Explanation>
SD0012	AESTDY is after AEENDY	Error	AE	5 (0.28%)	<Explanation>
SD0013	AESTDTC is after AEENDTC	Error	AE	1 (< 0.1%)	<Explanation>
SD0015	Negative value for SUDUR	Error	SU	3997 (96.76%)	<Explanation>
SD0017	Invalid value for RPTEST variable	Error	RP	53 (25.24%)	<Explanation>
SD0019	Invalid value for TSPARM variable	Error	TS	1 (3.70%)	<Explanation>
SD0021	Missing End Time-Point value	Warning	AE	576 (24.23%)	Subject XYZ-0001 did not have end dates or references collected and so could not populated in SDTM

Figure 8. Example of a Reviewer's Guide generated with Pinnacle 21 Enterprise

CONCLUSION



Validation is a process and the individuals tasked with creating and interpreting the validation report need to be organized and use the tools at their disposal to ensure an efficient end-to-end process. SAS macros can be utilized to make some process tasks easier; users also have technologies like JIRA to create and manage issues. Pinnacle 21 Enterprise offers several features specifically designed for issue management such as standard issue explanations and the ability to assign a validation issue to a specific individual. Regardless of the technology that is used, a successful validation issue manager knows the importance of validating early, tracking the validation results, and communicating findings to the right individuals.

REFERENCES

Amy Garrett and Chris Whalen. The Devil is in the Details – Reporting from Pinnacle 21 (OpenCDISC) Validation Report. *Proceedings of the PharmaSUG 2015 Conference*.

Ajay Gupta. Common Pinnacle 21 Report Issues: Shall we Document or Fix. *Proceedings of the PharmaSUG China 2018 Conference*.

Kristin Kelly. Best Practice for Explaining Validation Results in the Study Data. *Proceedings of the PharmaSUG 2018 Conference*.

RECOMMENDED READING

Sergiy Sirichenko. Common Programming Errors in CDISC Data. *Proceedings of the PharmaSUG 2017 Conference*.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Amy Garrett
Pinnacle 21, LLC
1777 Sentry Parkway W Building 17, Suite 405
Blue Bell, PA

Phone: 614-505-9043
Email: agarrett@pinnacle21.com
Web: www.pinnacle21.com

Brand and product names are trademarks of their respective companies.