

Introduction of check tool for SEND datasets via check list prepared by CDISC Japan User Group (CJUG) SEND team

Yuta Sakakibara^{1,2}, Atushi Hamada¹, Dai Yamamoto^{1,3}, Gen Sato^{1,4}, Junya Morita^{1,5}, Koichi Sato^{1,6}, Koji Akiba^{1,7}, Misa Matsumura^{1,8}, Sayaka Takami^{1,9}, Takashi Uota^{1,10}, Takayuki Yasuda^{1,10}, Terukazu Kitakahara^{1,11}, Yasuhiro Masuzaki¹, Yoshifumi Kaneko^{1,12}, Yoshinori Fujimura^{1,13}
1. CDISC Japan User Group, SEND team, 2. Kyowa Hakko Kirin Co., Ltd, 3. LSI Medience Corporation, 4. Eisai Co., Ltd., 5. Taisho Pharmaceutical, Co., Ltd., 6. Takumi Information Technology Inc., 7. Mitsui E&S Systems Research Inc., 8. SHIONOGI & CO., LTD., 9. Nihon Bioresearch Inc., 10. CAC Croit Corporation, 11. Instem Japan K.K., 12. KYORIN Pharmaceutical Co., Ltd., 13. FUJITSU LIMITED

1. Introduction:

When SEND datasets are created, we are expected to make the best efforts for ensuring their qualities based on many documents issued by the FDA and CDISC ^{1),2),3)}, but may meet SEND-specific difficulties, as follows;

- 1.Manual checks procedure is required to detect inconsistencies between the final report and SEND datasets, due to the different data creation flow of SEND datasets from that of TLFs of final reports through LIMS⁴⁾.
- 2. Variations among organizations in the internal rules for populating the values into SEND format variables and in their qualities⁴⁾.
- 3. The shortage of the resources with programming and data-handling skills in the non-clinical research field⁴⁾.
- 4. The de facto standard validator checks many rules, however it does not detect all issues.

Considering this situation, we created and published ([CDISC Portal > CDISC User groups > Asia > Japan > Open to Public > SEND > 05_CJUG SEND Check Tool](#)) a new list of common checks that each organization would like to confirm based on the above documents, and are also creating a SAS-based check tool to check SEND datasets automatically.

2. Check Levels and Target Range of CJUG SEND Team

We consider there are 5 levels of check procedures and their target qualities of SEND datasets as shown in Table 1. Levels 2 and 3 are the target ranges we think it important to achieve. Our check list and tool proposed here is designed to add additional check items in compliant with level 4 and 5 as an option.

Table 1 Descriptions of Check Levels

Level	Classification	Descriptions
1	Structural Check	Check by the de facto standard validator. Minimum requirement for SEND datasets submission
2	Essential Check	Check on Conformance to FDA and CDISC documents. Traceability to protocol and study report. Essential requirement for SEND datasets submission. CJUG SEND team added check items
3	Check across Organizations	Check according to a principle which can be shared across multiple organizations and which could contribute to decrease the difference in formats for populating SEND data across the organizations. CJUG SEND team added check items
4	Check across Studies	Check according to a principle which can be shared across multiple studies and which could contribute to better analysis after merging data from multiple studies in each organization. Each organization can add check items into CJUG SEND check tool.
5	Study Specific Check	Check in a specific condition in creating a SEND data for a specific study in each organization. Each organization can add check items into CJUG SEND check tool.

3. Method for preparing our check list

We jointly listed up check items referring documents issued by the FDA and CDISC. Knowledge from the team members’ experiences on QC of SEND datasets was discussed, and some of them were adopted as check items in the list. These items are listed in an Excel table for each domain which consists of maximum 8 columns (Table 2). We are currently creating check list in compliance with SENDIG v3.1 as next version.

Table 2 Columns of the check list

#	Column Name	Descriptions
1	Check Code No.	An unique code for each check item. The code corresponding to a SAS program of the check tool.
2	Variable Checked	Variable name of which the value is checked
3	Value of --PARMCD (For TS/TX only)	The TS.TSVAL of TS.TSPARAMCD or TX.TXVAL of TX.TXPARAMCD which is checked
4	Value of --PARM (For TS/TX only)	The TS.TSVAL of TS.TSPARAM or TX.TXVAL of TX.TXPARAM which is checked.
5	Check Algorithm	Description on the check algorithm
6	Check Importance	Two types of importance: Required or Recommended Required: Conformance to FDA and CDISC documents Recommended: Recommendations in FDA or CDISC documents, or Recommended format of CJUG SEND for populating the value
7	Check Needed Direct Review	If "Yes", directly reviewing the value is required. A list of values needed to review is exported.
8	Error Message	Error message text exported.

4. Structure of the check tool

Fig.1: Structure of the check tool

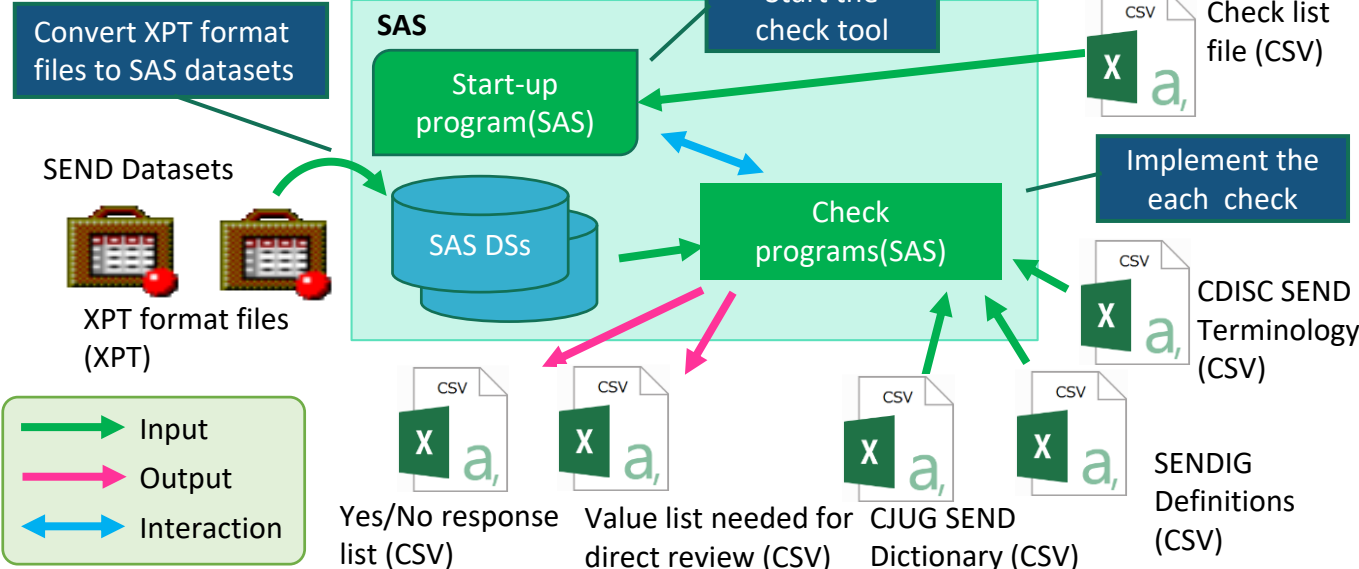


Table 3: Information files referred by programs

#	File Name	Descriptions
1	SENDIG Definitions	File included SENDIG definition (Ex. Domain Name, Variable Name)
2	CDISC SEND Terminology	CDISC SEND Terminology issued by NIH.
3	CJUG SEND Dictionary	A glossary that CJUG SEND team summarizes as reference dictionaries.(Ex. Company name of the member)

5. Discussion and Conclusion

CJUG SEND team members jointly created a new list of common checks and also we are making a SAS-based check tool to check SEND datasets. We would like these check list and check tool to be open to public in near future, so that each organization could incorporate them to their systems for checking the SEND datasets qualities. Using our check list and tool, we are expecting to reduce the risk in SEND data submission through the quality improvement, via performing the check which cannot be conducted in the existing validator. Also, it is expected that they can conduct unified checks across multiple organizations and can contribute to promoting efficient secondary use of SEND data, such as analysis of combined data from multiple studies, via reducing variations in SEND data. The check results output from the check tool would enable us to reduce workload on directly reviewing the values for check. Furthermore, using them, it would be possible to check the SEND data programmatically even by a person who do not have enough skills at data-handling. We believe that these check list and tool can contribute to solve the problems during SEND data preparation in Japan.

References

1)Technical Specifications Document; STUDY DATA TECHNICAL CONFORMANCE GUIDE. FDA., 2)Business Rules/Validator Rules, FDA, 3)Standard for Exchange of Nonclinical Data Implementation Guide (SENDIG). CDISC SEND Team, 4)Sakakibara Y. Collaboration across clinical and nonclinical team members to implement CDISC standards. PhUSE Single Day Event, Apr 14, 2017.

Abbreviations

CSV: Comma Separated Values, DSs: Datasets, LIMS: Laboratory Information Management System, SAS: Statistical Analysis System, TCG: Study Data Technical Conformance Guide, TLFs: Tables, Lists and Figures, XPT: SAS XPORT