Challenge for Dramatic Streamlining of TLFs Creation Process with SAS and Java

Yohei Takanami and Fumihiro Yamasaki, Takeda Pharmaceutical Co., Ltd.

ABSTRACT

It requires a large amount of time and effort to create tables, listings and figures (TLFs) and conduct QC/Verification of them in clinical trials. In this paper, we introduce a GUI-based system with SAS® and Java that enables users to generate over 90% of TLFs in typical clinical trials. The system has been successfully implemented in the actual TLFs creation process in Takeda in Japan, which has resulted in the reduction of workload, time and cost with keeping high quality. Key factors and considerations for streamlining TLFs creation process are also discussed.

INTRODUCTION

As mentioned above, TLFs creation process is complicated and it takes much time to complete the workload. In order to reduce the workload for TLFs creation efficiently, we had discussed the possibility of dramatic streamlining of business process. As shown in Figure 1, we concluded that a GUI-based statistical analysis system would be one of the best solutions to achieve multiple benefits in our business process.
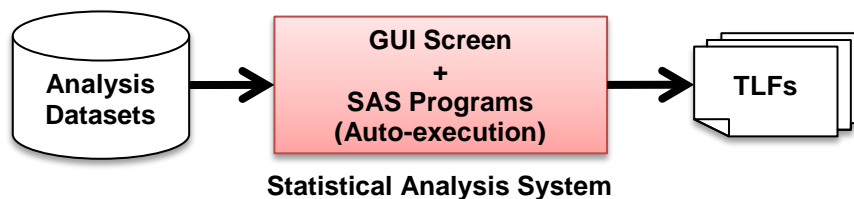

Figure 1. Concept of statistical analysis system

Before moving to the discussion of system development, there were many issues to be overcome in actual operations of TLFs creation in clinical trials. The following are major issues and considerations on system development and maintenance.

- Understanding of benefits of utilizing a statistical analysis system among stakeholders

- Standardization of TLFs layout, statistical methods and analysis data structure

- Resources for the system development and maintenance for long-term use

After the discussion above, we have developed SAS macro library and GUI screen with Java. The reason why we have used SAS and Java are as follows:

- Needless to say, SAS is a gold standard in pharmaceutical industry, therefore most statisticians and programmers are familiar with SAS. In order to develop SAS macros efficiently, skilled SAS programmers need to be assigned to the development team.

- Java is widely used to develop applications regardless of fields. Its useful functionalities can be applied to a system development in combination with SAS. We can also update easily the version of Java without dependence on other software products.

Finally, we decided on architecture of the system as shown in Figure 2. Those who utilize the system log in to the system using authorized ID and password via Internet Explorer (IE). Java stores and passes input parameters and arguments in XML format and then executes SAS macros with SAS Integration Technologies (IT). SAS macros convert parameters passed by Java to SAS macro parameters and generate reports using ODS RTF or ODS TAGSETS.RTF statements.

- SAS server (Unix): SAS 9.4 and SAS IT
  SAS macros, input parameters, analysis datasets and analysis results are stored in this server.

- Application server (Windows): Java (Java Development Kit (JDK) 1.6.0_39) with Tomcat (7.0.37)
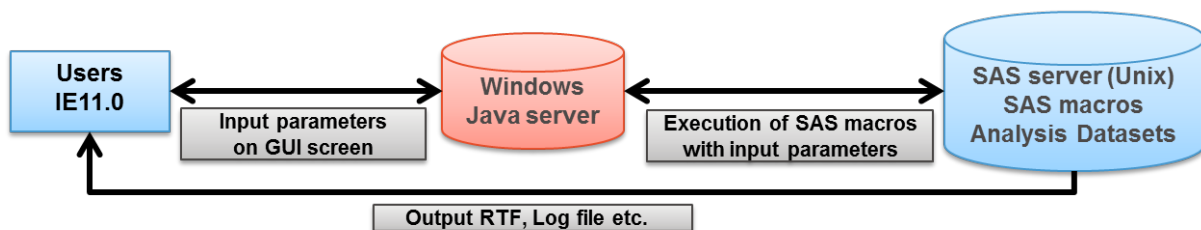 This server hosts all the Java server-side components.

Figure 2. System architecture

Details of implementations and functionalities of the system are discussed in the next section.

IMPLEMENTATION OF GUI-BASED SYSTEM INTO BUSINESS PROCESS

In this section, we introduce the overview of the system development, major functionalities and benefits of business model with the statistical analysis system.

SYSTEM DEVELOPMENT

As mentioned in the previous section, we had a lot of discussion about business model and standardization of statistical deliverables before the development phase. The key materials and items are summarized in Table 1. Especially standardization of TLFs and analysis data structure among stakeholders (e.g. clinical scientist, statisticians, SAS programmers, medical writers) are keys to development of the system and streamlined process.

Materials/Items	Discussions/Recommendations
Standard TLFs	Standard templates were discussed with stakeholders in each region. Further, typical TLFs (including options) used in clinical trials were developed and maintained in Takeda. Standard TLFs consist of over 140 templates.
Statistical Methods	Typical methods used in clinical trials (including options) were discussed with stakeholders in each region. Japanese specific methods were added to the local statistical test library.
Analysis Data Structure	Takeda standard structure has been shifted to the Clinical Data Interchange Standards Consortium (CDISC) Analysis Data Model (ADaM) in recent years.
Operators	Those who are responsible for input parameters based on Statistical Analysis Plan (SAP), dataset specification and so forth. Operators are unnecessarily familiar with SAS programming and statistical analysis.
SAS modules	Internal SAS programmers have developed and maintained SAS modules in terms of cost and maintainability.
Java modules	There are few Java programmers in pharmaceutical companies, therefore external resources should be used to develop Java modules.
Project management	Appropriate IT people should lead and manage the project to meet the Computerized System Validation (CSV) requirements and so forth.

Table 1. Summary of key materials and items to system development

After the materials mentioned above were agreed and finalized among stakeholders, system development was conducted by the team in accordance with company's standard procedure of system development and CSV policy as shown in Figure 3. We have been utilizing this system in actual workload in Takeda in Japan and maintaining with the same policy since its release in 2009.
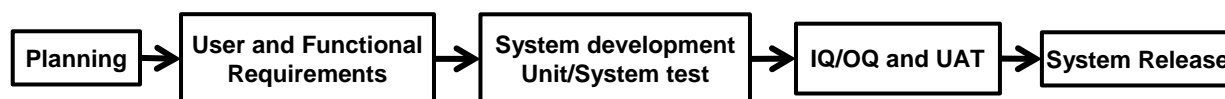
| Planning | → | User and Functional Requirements | → | System development Unit/System test | → | IQ/OQ and UAT | → | System Release |

Figure 3. System development flow

SYSTEM FUNCTIONALITIES

The system has over 140 TLFs and 50 statistical test macros. In addition to them, there are also utility macros such as metadata management, common logic and so forth. Table 2 and Table 3 show the examples of templates and statistical test/method macros implemented in the system respectively.

Categories	Types of templates
Disposition etc.	Disposition, Baseline and demographic characteristics, Medical History etc.
Efficacy	Typical analyses for continuous variables and binary/categorical data, Survival analysis, ANCOVA, Logistic regression model, Cox regression model, non-inferiority test etc.
PK	Summary of PK concentration and PK parameters, ANOVA for BE studies etc.
Safety	Adverse events, Laboratory test results, Vital signs, ECG parameters etc.
Figures	Plot per subject, Mean plot, Bar chart, Box plot, Scatter plot etc.
Listings	Typical listings in CSR Appendix

Table 2. Examples of TLFs templates implemented in the system

Categories	Statistical tests/methods
Confidence interval	Mean, Proportion, Mean difference, Risk difference etc.
Continuous variables	t-test, ANOVA, ANCOVA, non-inferiority test etc.
Binary/Categorical data	Fisher exact test, Chi-square test, non-inferiority test, MH test, CMH test etc.
Multiple comparison	Bonferroni, Dunnet, Tukey, Williams etc.

Table 3. Examples of Statistical test macros of typical methods implemented in the system

Logon and Parameter Input Screen are shown in Figure 4. The users log on to the system via the logon screen and then select a study and a template.
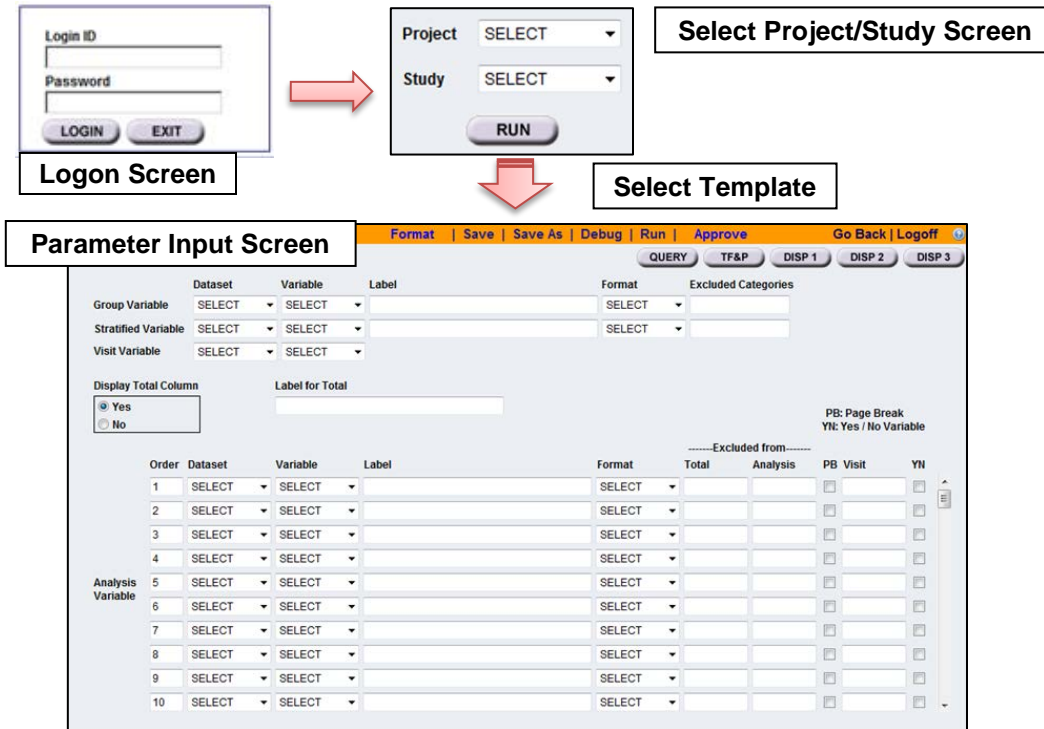
Figure 4. Logon and Parameter Input Screen

Once a parameter input is completed on a screen, the parameters are stored in XML format in SAS server by clicking on "SAVE" button and then SAS macros are executed by clicking on "RUN" button. Figure 5 shows the execution flow of Java and SAS modules for one TLF creation. Finally, the system outputs a RTF and a log file in the designated folder in the SAS server.
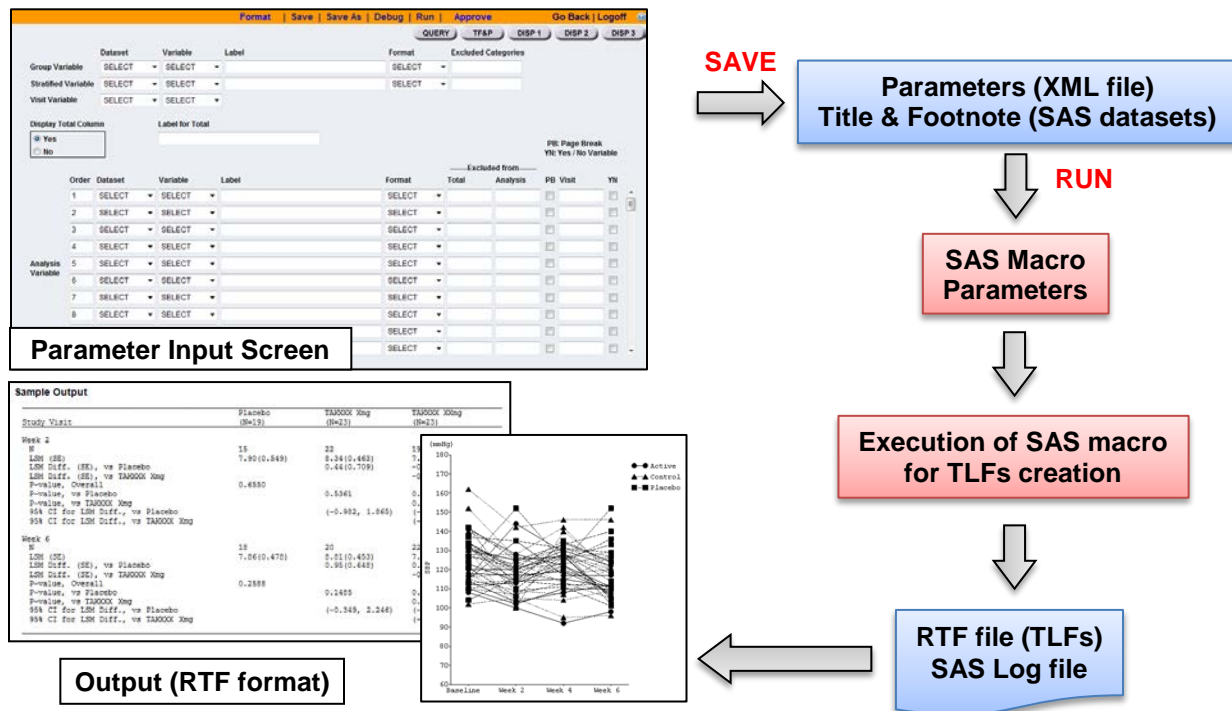

Figure 5. TLF creation flow

Table 4 shows an image of XML parameters and converted SAS macro parameters. After the "RUN" button is clicked, the system converts the XML parameters into "%LET" statements which are used in a TLFs SAS macro.

XML parameters	SAS macro parameters

`<template_name>TEST</template_name>`
  `<D_M_GDS>ADSL</D_M_GDS>`
  `<V_M_G>TRTSEQP</V_M_G>`
  `<_M_GF>$200.</_M_GF>`
  `<Analysis_Variable>`
    `<_ORDER>1.0</_ORDER>`
    `<D_ITMDS>ADSL</D_ITMDS>`
    `<V_ITMV>AGE</V_ITMV>`
    `<_ITML>Age</_ITML>`
    `<_ITMF>8.0</_ITMF>`
	`%LET _ITMV1 = ADSL|AGE ;`
	`%LET _M_G   = ADSL|TRTSEQP ;`
	`%LET _M_GF  = $200. ;`
	`%LET _ITML1 = Age ;`
	`%LET _ITMF1 = 8.0 ;`

Table 4. Image of XML and SAS macro parameters

Figure 6 shows the main process of SAS macros for TLFs creation after receiving input parameters from Java side. The flow varies according to template type (e.g. Table/Graph, Disposition/Safety/Efficacy) in detail.
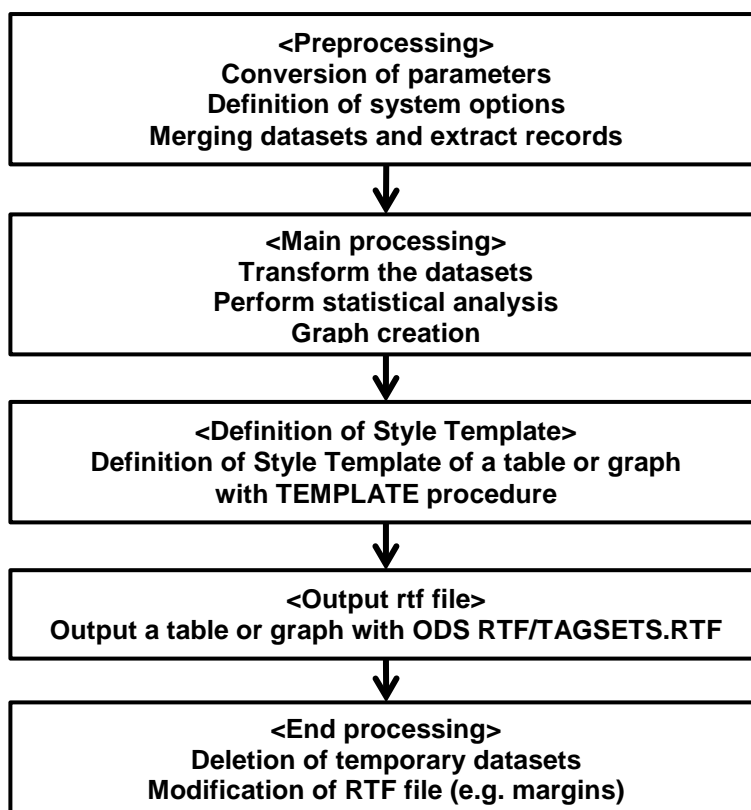

Figure 6. Processing flow of TLFs creation macro

The system can also generate more than one TLF at a time with Batch Job functionality. In this functionality, a user can select templates from the list and then execute all the specified SAS macros with "RUN" button. This functionality is utilized after completion of TLFs parameters in each template.
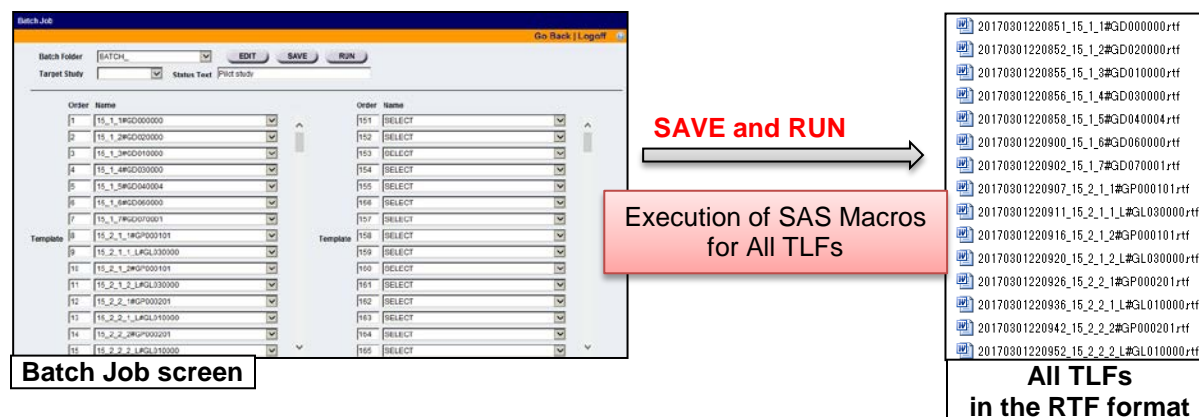
Figure 7. TLGs creation with Batch Job functionality

BUSINESS PROCESS AND BENEFITS WITH THE STATISTICAL ANALYSIS SYSTEM

As shown in Figure 8, we have developed and implemented the statistical analysis system for TLFs creation and realized the streamlined process. Once analysis datasets are ready, system administrators set up them and then operators can input parameters in accordance with SAP, specifications of analysis datasets, standardized TLFs templates and relevant materials. Statisticians and programmers give explanations to operators about input parameters whenever necessary.
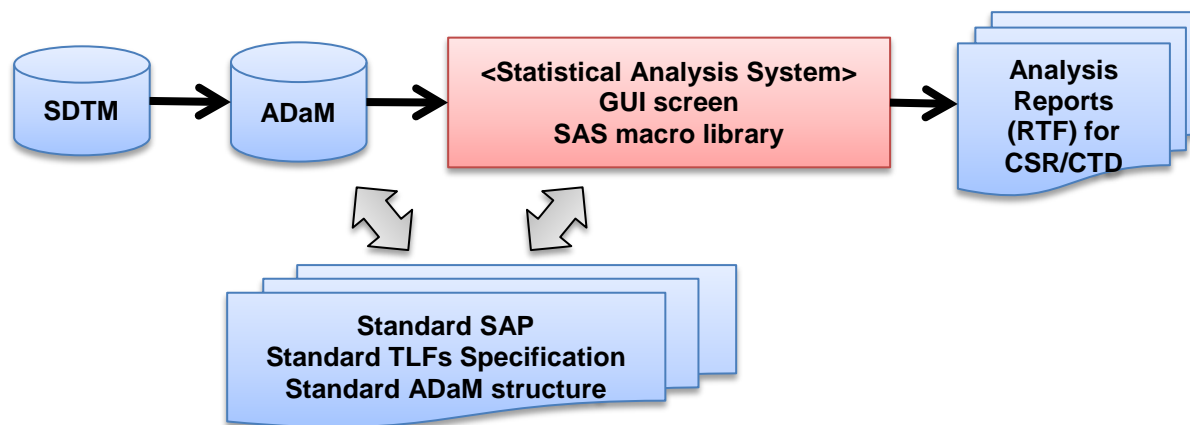

Figure 8. Business flow with the Statistical Analysis System

As a result of implementation of the statistical analysis system, we achieved multiple benefits as follows:

- Reduction of workload, time and cost, while keeping high quality

  - Over 90% of TLFs can be generated by the system with typical clinical trial design and analysis. For non-standard templates which are not implemented in the system, although the users have to create SAS programs manually, these programs can be executed in the system.

  - When compared with outsourced model, costs of in-house model with the system can be reduced as much as 50% or more for TLFs creation.

  - Maintenance and license costs are not so high (e.g. Server SAS, system revision).

  - High maintainability was realized due to the simple architecture with SAS and Java.

- Flexibility for analysis dataset structure

  - Takeda legacy datasets and CDISC ADaM structure can be utilized in the system.

- Skills required to operators

    ➢ Operators are unnecessarily familiar with SAS programming and statistics for input parameters (need the SAS programming skills for the creation of non-standard programs).

- Disk space saving

    ➢ Parameter files are saved in XML format, therefore the file size can be reduced compared with other non-text file formats.

- Conformance to requirements of the 21 CFR Part 11

    ➢ CSV was conducted and the system meets the requirements of the 21 CFR Part 11 (e.g. user access, audit trail).

DISCUSSIONS

As shown in the previous section, we have achieved multiple benefits on our business process with the statistical analysis system. However, there are several considerations to utilize the system in the actual workload efficiently as follows:

- Standardization of statistical deliverables (e.g. statistical methods, TLFs templates, analysis data structure)

    ➢ System requirements vary according to region, therapeutic area, organization etc., therefore it is very important to discuss harmonization of statistical deliverables among stakeholders to the maximum extent possible prior to system development.

    ➢ Some additional variables specific to the system requirements need to be included in the analysis datasets.

- Possibility for global or multiregional use

    ➢ For global pharmaceutical companies, it is better to use the same system, environment and business model in terms of process harmonization. In order to realize the global or multiregional use, we have to address differences of organization and requirements between regions.

    ➢ System performance and maintenance window for planned downtime (e.g. backup, installation of security patch) are also important issues due to time difference and operation at different locations.

- Development of Business process

    ➢ Need to discuss which is better to use In-house or Outsourced model based on organization and role and responsibilities in each company.

- System users

    ➢ Need to discuss who should be users/operators of the system. SAS programming skill is required for the creation of non-standard programs.

- Electronic Data (e-Data) submission for New Drug Application (NDA)

    ➢ Global pharmaceutical companies have to address the requirements on e-Data submission to regulatory agencies (e.g. FDA, PMDA).

    ➢ Although the regulatory agencies may request sponsors to submit the analysis programs, the executed programs cannot be extracted from the current version of the system. PMDA strongly recommends sponsors to submit Analysis Results Metadata (ARM) which includes analysis programs.

CONCLUSION

GUI-based statistical analysis system to generate TLFs can be a very useful tool in terms of reduction of workload, time and cost in clinical trials. Although the discussions toward the realization of global use are underway, the system introduced in this paper has been implemented successfully in Takeda in Japan. Standardization of statistical deliverables (e.g. statistical methods, TLFs templates, analysis data structure) is the most important factor and should be discussed among stakeholders prior to the system development especially for global use.

REFERENCES

Mo, Albert. 2006. "Developing, Managing, and Evaluating a Standard Macro System." *Proceedings of the 31th Annual SAS Users Group International Conference.*

Grandits, Greg. 2006. "A System of SAS® Macros for Producing Statistical Reports." *Proceedings of the 31th Annual SAS Users Group International Conference.*

ACKNOWLEDGMENTS

We would like to thank Takamasa Hashimoto, Koji Iwamoto and Yasunori Endo for sharing their experiences and expertise.

RECOMMENDED READING

- *ICH Harmonised Tripartite Guideline: Structure and content of clinical study reports E3. Step 4. 1995.* (http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E3/E3_Guideline.pdf)

- *ELECTRONIC CODE OF FEDERAL REGULATIONS TITLE 21 PART 11* (http://www.ecfr.gov/cgi-bin/text-idx?SID=459fd12676f041a5be64496e2cae342f&mc=true&node=pt21.1.11&rgn=div5)

- *CDISC Analysis Data Model (ADaM)* (http://www.cdisc.org/adam)

- *FDA Study Data Standards Resources* (www.fda.gov/ForIndustry/DataStandards/StudyDataStandards)

- *PMDA Advanced Review with Electronic Data Promotion Group* (www.pmda.go.jp/english/review-services/reviews/advanced-efforts/0002.html)

CONTACT INFORMATION

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

Yohei Takanami
Takeda Pharmaceutical Co., Ltd.
youhei.takanami@takeda.com

Fumihiro Yamasaki
Takeda Pharmaceutical Co., Ltd.
fumihiro.yamasaki@takeda.com

SAS® and all other SAS® Institute Inc. product or service names are registered trademarks or trademarks of SAS® Institute Inc. in the USA and other countries. ® indicates USA registration.