PharmaSUG China 2017 - Paper 67

Overview of OSI Deliverables from Programming Perspective

Wang Zhang, MSD R&D (China) Co. Ltd., Beijing, China

ABSTRACT

The Office of Scientific Investigation (OSI) at U.S. FDA is responsible for the verification of the integrity of the clinical efficacy and safety data. It requests the applicants of NDA(s)/BLA(s) and their supplemental filings to submit summary level clinical site data, line listings by site and other deliverables, and it would be able to use specific algorithms to identify high risk sites and conduct site inspections. This paper offers a brief introduction to the contents of these OSI deliverables, and highlights some take-home messages in the preparation of the package from the statistical programming's point of view.

Key Words: Office of Scientific Investigation, OSI Deliverables

INTRODUCTION

Office of Scientific Investigation, or known as OSI, is now under the Office of Compliance, in the Center for Drug Evaluation and Research (CDER), FDA. With the objectives of verifying the integrity of efficacy and safety data submitted to the FDA for new drug applications, and more importantly, protecting the rights and welfare of human research subjects, OSI is focusing on the audit and verification on the clinical trial data, the directing of inspections of Institutional Research Boards (IRBs) for compliance with standard, and ensuring the non-clinical and clinical studies for new drug development are conducted in the compliance with United States laws and regulations[1].

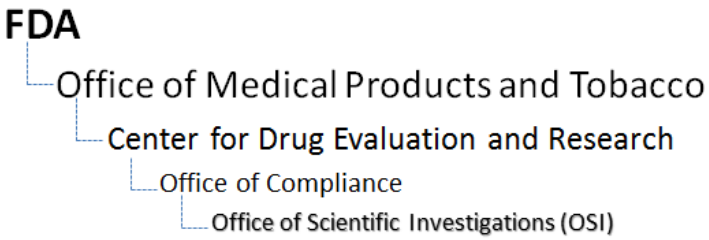

Figure 1. OSI Organization Chart in FDA

OSI will work with other FDA departments and offices to conduct inspections on sponsors, CROs and IRBs to ensuring the clinical trial data integrity, and only when the offices are able to access the adequate and accurate data during the pre-approval inspection planning phase, could they have the onsite inspections in a more timely and efficient manner. In October, 2012, OSI hosted a webinar 'Overview of Information Requested by CDER's Office of Scientific Investigation (OSI) for NDAs and BLA submissions' and addressed three parts of deliverables that it requested from applicants or sponsors to assist the inspections[2]. Subsequently, in December of the same year, CDER released a draft and more specific guidance for industry, with regards to the content and electronic submission requirements of the 'Summary Level Clinical Site Data', which served as the 3rd part of the OSI deliverables[3]. The OSI package includes:

- Part I Tabular Listings of Site Information
- Part II Line Listing by Site
- Part III Site-Level Dataset (Voluntary but recommended)

And they will be submitted to the agency along with other NDA(s)/BLA(s) and their supplemental applications submission packages. The requirements of each part will be described in below sections, respectively.

For the submissions to the agencies or departments other than FDA CDER, this package is out of scope.

PART I TABULAR LISTINGS OF SITE INFORMATION

This part actually involves multiple deliverables, including site information, study original protocol and all its amendments, and the sample annotated case report forms. The sites participated or referenced in all the studies submitted for the U.S. NDA(s)/BLA(s) and their supplemental applications, and the documents for all these studies, should be covered in this part. All these items should be in PDF format, and the contents of deliverable items are summarized as below:

Item	Content	Reason for Request	Comment
1	Site Information, inclusive of site number, name, address and contact information (Phone, Fax, Email)	Site inspections would need to be conducted at the correct location, with accurate contact information.	Updated site contact information should be submitted if the contacts change.
2	High level study summary by site, such as the numbers of subjects screened, enrolled, discontinued and with protocol violations	This could support to facilitate a timely site selection and inspection assignments.	
3	A list of responsible entity, to which the sponsor contracted the key study functions, e.g. drug distribution, data management; And also the location of study-related documents, e.g. data management plan, training records and so on.	This will provide the correct and responsible party for the inspection, and the right area of study files for agency to review.	For pivotal studies.
4,5	Study original protocol and all the amendments, study annotated case report form	They are for FDA investigator to evaluate the protocol adherence.	These items do not need to be re-submitted if they present in the eCTD package elsewhere, applicants could provide the hyperlink to them in the reviewer's guide.

Table 1. OSI Part I Deliverable Contents

In practice, item 1 and 2 might be combined as the following example:

Site	Number of Subjects Screened	Number of Subjects Enrolled/Randomized	Number of Subjects Discontinued	Number of Subjects with Protocol Violations
Site No. ## Jim Investigator, M.D. Study Center, 666 Research Road, Trial City Phone: 123-456-7890 Fax: 123-456-7890 Email: Jim.Inv@center.com	50	45	5	10

Table 2. Site Information Listing Example

PART II LINE LISTING BY SITE

This part involves a series of 'line' listings by site in PDF format, which contain most of the data fields of the study and may satisfy the FDA RA department's needs to verify the data. These listings are quite like CSR listings, however, the major difference between the line listings and CSR listings is, the line listings should only adopt preliminary/raw data points but no derived data or endpoints, it is quite essential for the data verification. Plus, the CSR listings are usually sorted by treatment arm and subject ID, but the line listings are presented site by site. The office has outlined 10 listings and their related contents as to be requested, but applicants are encouraged to propose any alternative materials or layouts of the listings, if they could benefit the inspection and review, during the pre-approval meeting discussion with FDA. The mandatory types of listing as requested are:

1. Listing for each subject/number screened and reason for subjects who did not meet eligibility requirements
2. Subject listing for treatment assignment (randomization)
3. Subject listing of drop-outs and subjects that discontinued with date and reason
4. Evaluable subjects/ non-evaluable subjects and reason not evaluable
5. By subject listing of eligibility determination (i.e., inclusion and exclusion criteria)
6. By subject listing, of AEs, SAEs, deaths and dates
7. By subject listing of protocol violations and/or deviations reported in the NDA, description of the deviation/violation
8. By subject listing of the primary and secondary endpoint efficacy parameters or events. For derived or calculated endpoints, provide the raw data listings used to generate the derived/calculated endpoint.
9. By subject listing of concomitant medications (as appropriate to the pivotal clinical trials)
10. By subject listing, of laboratory tests performed for safety monitoring

And it would be preferred that the listings could be categorized and filed in this structure:

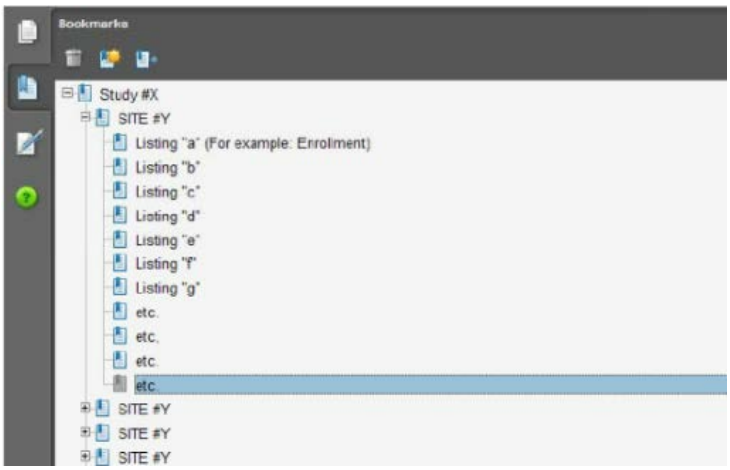

Figure 2. Line Listing by Site

As for the SAS programming practice in our company, we incline to create subfolders for each site involved first, and generate each type of line listings by site with do-loop statement into the folders, eventually combine them into one file (or split it into several files afterwards due to the file size limitation). Moreover, we prefer to add the site ID, country of site and investigator names into the header of the listings for reviewer's convenience:

```
proc sql noprint;
  create table site as select distinct studyid, sitenum, invnam,
cntrynam,strip(studyid)||'-'||strip(sitenum) as study_site from adsl;
  select count(distinct study_site) into: totsite from site;
quit;
%let totsite=&totsite.;
```

```
proc sql noprint;
  select distinct study_site into: site1-:site&totsite from site order by study_site;
quit;

%do j=1 %to &totsite;

  %*Create subfolder for each site *;
  Data _null_ ;
    RC = DCreate( "site&&site&j" , path) ;
    If RC ne " " then put RC= ;
  Run;

  %*Create listings for each site *;
  Proc report data=xxxx;
    Xxxxxxxxxxx;
  Run;

%end;
```

PART III SITE LEVEL SUMMARY DATASET

In order to maintain a prompt approach for site selection, FDA CDER is committed to develop a clinical site selection tool from a risk-based model to identify the clinical sites with high risk more efficiently. This tool would rely on a model that takes an array of risk parameters associated with the sites and provides targeted sites as the results of certain algorithm. Therefore, the site level summary dataset, which should consist of all the pivotal studies and sites involved in the application, and the summarized characteristics and outcomes of each individual study site, is a key input element to this tool and model, and it is optional for sponsors to generate this dataset and have it submitted as a supportive file. Meanwhile, sponsors could take the advantage of this dataset and run the model by themselves to identify any sites with potential issues in advance, for a better inspection preparation.

It is allowed that the dataset contains multiple pivotal studies participated in the application. If the applicants agree to submit this dataset during the discussion with FDA, the data records should be summarized and presented by study, by site and treatment arm with the population that supports the efficacy analysis of the studies, so that one site might go across several data records depending on the number of studies and treatment arms it holds. The dataset should be named as 'clinsite.xpt' in the XPT file format, and a DEFINE file in the PDF format would be preferred to be submitted as well. The dataset sample specification drafted and released by CDER could be taken as a reference[4]. To elaborate on the data format and supportive file, the XPT, or SAS Transport Format (XPORT), is the file format for the submission of all the electronic clinical datasets requested and accepted by the FDA. It's an open file format published by SAS Institute, and it could be transformed to other commonly used file format without inevitably depending on SAS program or other specific vendor, which satisfies FDA's 'vendor neutral' principle. XPT files could be generated using COPY Procedure in SAS version 5 and above as FDA would accept. Please do not use SAS CPORT to create the XPT files since FDA cannot review or process such type of generation[5]. Some sample codes to create the XPT files are as below:

```
%* Define the library where the SAS dataset(e.g. clinsite.sas7bdat) are stored;
libname source 'SAS-library-on-sending-host';

%*Define the library where the XPT files are to be stored;
libname xptout xport 'filename-on-sending-host';

%*method 1;
proc copy in=source out=xptout memtype=data;
  select clinsite;
run;
```

```
%*method 2;
data xptout.clinsite;
  set clinsite;
run;
```

And the DEFINE.PDF is a supportive document that demostrates the attributes of the submitted dataset, including the dataset name, label, description, and the names, types, lengths, codelists, sources and derivations of all the variables included. The document could directly provide the derivation details to the reviewer, and it is usually similar to the specification we were using to create the data. But the author should make efforts to use English language instead of programming codes in the derivations as possible as one could, for the agency reviewer's concern.

Furthermore, with regard to the clinical data in eCTD format, all the OSI deliverables, including part 1 to part 3, would need study tagging files (STF) for each study with the appropriate leaf titles. The part 1 and 2 deliverables will locate in each study folder, and part 3 dataset could rest in m5 folder during the submission, until CDER determines and rolls out the formal folder structure for the OSI package. A reviewer's guide in PDF format, which is not mandatory and may include the location and link to each component of the package, could also be submitted.

DISCUSSION AND LONG TERM VISION

The statistical programmers from sponsors or CROs might take more responsibilities in preparation of the OSI deliverables, but it absolutely requires the collaboration with other functions have the package complete. For example, in part 3 site level summary dataset, the statistics and deviations for primary efficacy endpoint should be summarized by treatment and by site, whereas this calculation might not always occur in the rest of the submission such as CSR reports. Hence, programmer should check and work with the statistician to request and determine the derivation rules for these variables. It's also important to confirm with the statistician on the population that the summary is extracted from, and all the variable flags that the summary would use. Besides, all the 3 parts of deliverables might require the data beyond the clinical database, like site contact information, study administrative IDs or protocol violations. Programmer should follow up with other team members such as clinical or regulatory affairs on these external data as well.

It would be suggested to do the cross check among the OSI deliverables before the delivery. For example, the programmer should compare the summarized number of subjects with serious adverse events in the site level summary dataset, with the counts of subjects with events marked as serious in the AE line listings, for the same and single site. Programmers are also encouraged to check the content in the OSI deliverables against the CSR report in case any discrepancies, the development of tools for these cross checks would be preferred.

So far the CDER has not requested the applicants to submit the source data for the OSI deliverables generation, it has not determined the data standard followed by the site level summary dataset either, however, it is the long term objective for CDER and the offices to develop and implement the data standard such as CDISC, the electronic data submission and access, and the more effective and user-friendly review tools specifically for the OSI package. Industry stakeholders are encouraged to provide inputs and thoughts to the agency with regards to these developments. Among the deliverables, part 1 and part 2 are required as the mandatory components of the submissions to CDER after October, 2012, the absence of these deliverables from the package may cause the delay of agency review and approval timeline. Part 3 dataset is a voluntary submission, which would support the office for site selection, and the applicants may consider to submit it if the preparation would not impact the timeline. Above all, the discussion and confirmation between the agency and applicants on the contents, the preparation and the submission timeline of the OSI package during the pre-appoval process is very necessary.

CONCLUSION

OSI deliverables are critical parts of the NDA(s)/BLA(s) and their supplemental applications submission package to FDA CDER. Statistical programmer would take responsibilities in generating and delivering the package according to the guidelines from the agency, but it's also a collaborative activity with other team members involved. CDER is making efforts to develop the submission standard and review tools for the OSI package in the pursuit of conducting more efficient and timely inspections, therefore as the applicants, industry sponsors are welcome to propose the innovative approaches and communicate with the agency, as well as keep track of the latest updates from the regulatory office.

REFERENCES

1. FDA CDER OSI Homepage

https://www.fda.gov/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDER/ucm090085.htm

2. Ann Meeker-O'connell, Sean Kassim. FDA Webinar: Overview of Information Requested by CDER's Office of Scientific Investigations (OSI) for NDAs and BLA submissions. December 2012. Available at

https://collaboration.fda.gov/p44198603/?launcher=false&fcsContent=true&pbMode=normal

3. U.S. FDA CDER. Guidance for Industry: Providing Submissions in Electronic Format – Summary Level Clinical Site Data for CDER's Inspection Planning, Draft Guidance. December 2012. Available at

https://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/FormsSubmissionRequirements/UCM332468.pdf

4. U.S. FDA CDER. Specifications for Preparing and Submitting Summary Level Clinical Site Data for CDER's Inspection Planning. November 7th, 2012. Available at

https://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/FormsSubmissionRequirements/ucm332466.pdf

5. U.S. FDA CBER and CDER. Study Data Technical Conformance Guide. March 2017. Available at

https://www.fda.gov/downloads/ForIndustry/DataStandards/StudyDataStandards/UCM384744.pdf

ACKNOWLEDGMENTS

The comments for this paper from Greg Zhou and Yanhong Li are appreciated.

RECOMMENDED READING

FDA Webinar: Overview of Information Requested by CDER's Office of Scientific Investigations (OSI) for NDAs and BLA submissions. December 2012.

CONTACT INFORMATION

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

Name: Wang Zhang
Enterprise: MSD R&D (China) Co. Ltd.
Address: Building 21 Rongda Road, Wangjing R&D Base, Zhongguancun Electronic West Zone, Chaoyang District
City, State ZIP: Beijing 100012
Work Phone: 86 10 58609407
E-mail: wang.zhang@merck.com