**Impacts of the CDISC Standards on our Clinical Study Operation: Lessons and Current Achievements**

Yoji Nagai, MD, PhD, on behalf of the CDISC project team at Translational Research Informatics Center

Translational Research Informatics Center, Foundation for Biomedical Research and Innovation, 1-5-4 Minatojima-minamimachi, Chuo-ku, Kobe, Hyogo 650-0047, Japan

TEL: +81-78-303-9093, FAX: +81-78-303-9094, E-mail: nagaiy@tri-kobe.org

**Q1. What is TRI center, and what kinds of activities does it have?**

Translational Research Informatics (TRI) center is one of the institutes that belong to the Foundation of Biomedical Research & Innovation in Japan. It was founded in 2003, given financial supports from the Japanese government and Kobe city. Particularly, TRI center is committed to the innovation of medical technology, construction of infrastructure and creation of new science, through the facilitation of translational and clinical researches. Also, it serves as the largest data center in Japanese academia, currently supporting 75 clinical studies of various kinds. To promote health and welfare of all people, we are open to all medical researchers and provide comprehensive study support services, including preclinical consultation, protocol development, data management, system development and statistical analysis.

**Q2. Why did you embark on the world of CDISC?**

Since the beginning in 2003, we have been supporting more than 160 clinical studies. However, the data formats were not standardized and we were dependent on several data management systems, imposing additional burdens to the staffs and precluding integration of the pooled data. Recently, in the face of increasing number of concurrent studies, we had practical needs to increase efficiency in the study operation, and to analyze combined data from relevant studies. Fortunately, we had kinds of opportunities, including skilled staffs and accumulated experiences. Thus, the circumstances surrounding us were maturated, and pressed by need and opportunity, we had decided to get into the world of CDISC.

**Q3. How are you utilizing the CDISC standards at TRI center?**

At the CDISC International Interchange in 2011, I made a presentation that we were applying CDASH/SDTM to an Alzheimer’s disease (AD) study, and were in the final stage of developing our original electronic data capturing (EDC) system. Thereafter, the AD study was started in January of 2012, and more than 130 subjects have already been recruited. Also, development of the EDC system was completed, with all regulatory requirements satisfied and all necessary computer system validation (CSV) performed.
Because of the nature of TRI center, most of our CDISC related activities have been limited to CDASH/SDTM. Initially, the activities were started with the preliminary application to several studies and extraction of the core processes. Through the execution of relevant work, a set of operation manuals was formulated, the processes of which served as on-the-job training for us. At current, we are working to refine the internal rule for assigning CDISC variables. Indeed, in addition to the above mentioned AD study, we have applied CDASH/SDTM to more than 10 studies of various diseases. The first 3 studies are being conducted on paper CRF, and thus we picked up the variables from CRF, and deployed them in the local data management system. Regarding the rest of newer studies, we deployed the variables in aforementioned EDC system, in accordance with our CDISC implementation guide and CRF design manual. As for the diseases for which therapeutic area standards have not been published, we made every effort to comply with the standards, but created our own variables by necessity, following the manner of CDASH/SDTM.

As for the new EDC system, which was named as “eClinical Base,” we are using it for all studies newly supported at TRI center. Most outstanding feature of the system is that we prepare the setting specification in Excel files, including all variables and parameters required for the setting. When a draft version of study protocol is ready, the most similar setting specification is selected from the file library of prior studies. Then, the selected setting specification is modified by clinical data managers to comply with the study protocol, in accordance with our CDISC implementation guide and CRF design manual. Subsequently, it is polished up by technical data managers, and then finalized by system engineers. When the specification is fixed, the Excel file is imported to the EDC system, whereby the data entry interface is automatically generated. On the other hand, all such Excel files are stored by disease categories, and they are reused when setting a new study. By repeating this cycle, the core sets of variables will be shaped for respective diseases, which increases efficiency of the study setup and system preparation. In fact, we have begun to experience a virtual increase of efficiency for certain diseases.

Q4. What have you learned through the application of standards to AD study?

The AD study, to which we have formally applied the standards for the first time, is a prospective cohort study, enrolling 400 subjects in Shanghai, China. The study purpose is to identify individuals with mild cognitive impairment, who convert to AD. I was responsible for the protocol development, with a dedicated team of neurologists, protocol managers, data managers, biostatisticians and system engineers. When developing the protocol, a series of neuropsychological tests was selected from the worksheets actually used in Shanghai, and the variables were defined by referring to CDASH/SDTM. However, we were faced with a couple of problems during the processes.

More specifically, we found 2 tests that differ between US and China, precluding direct applications of the standards. Those are ADAS-Cog and AVLT, where the former is a common scale to evaluate overall cognitive function, whereas the latter is a common test for words recalling. Thus, we solved this issue by creating new variables, rather than compulsively applying the predefined standard variables. In part due to such a flexible solution, data are smoothly entered in Shanghai, and no inconvenience has been reported so far.
From the experience with AD study, we have learned that there are kinds of obstacles when applying the standards to different cultures, such as to Japan or to China. Particularly, as for neuropsychological tests, there exist at least two levels of obstacles, which are apparent or intrinsic, sometimes requiring a scrutiny, but more often demanding deep scientific insights. More specifically, the first level of obstacle is the apparent or superficial differences in the questionnaires, scoring methods or data structure of the tests. Indeed, as mentioned above, ADAS-Cog is somewhat different by countries. For instance, it is called ADAS-RC in China, where “number cancellation” is replaced by “concentration”, compared to the US version. In Japan, ADAS-Cog in US is modified and called as ADAS-J Cog, where 13 items of questionnaire are reorganized into 11 items. Still, the maximum score is the same among 3 countries. These are the obstacles readily found through a scrutiny of respective neuropsychological tests. By contrast, the second level of obstacle is related to the intrinsic or essential difference in the linguistic or cultural background. Such obstacles can be only recognized through meticulous insights by experienced neurologists with different cultural backgrounds. For example, in case of MMSE, which is a common simple measure for cognitive function, the scoring method and maximum score are the same among 3 countries, but questions included in the test are slightly different. Then, how can we warrant that 2 points decrease in MMSE score imply the same cognitive decline in US, Japan and China?

These obstacles are most profound when comparing cognitive function across different cultures, or when performing integrated analysis of studies from different countries. Indeed, neuropsychological tests are influenced by many factors, including literacy rate, education years and wording of questionnaire. Thus, in addition to the standardization of data format and terminology, standardization of the questionnaire and intercultural validation may be required, to allow for the comparison or integration. This is a major scientific challenge imposed on both neurologists and clinical study specialists, which would be overcome only through a tight collaboration of motivated researchers from different cultures. Also, because it is virtually impossible to simultaneously standardize many kinds of neuropsychological tests, we may need to begin with the most important tests such as ADAS-Cog and MMSE.

Q5. How did the CDISC standards impact on your clinical study operation?

Yes, the CDISC standards have greatly and positively impacted on our process of clinical study operation. Let me briefly explain with regards to the notion that “CDISC is more than Standards”, which was found on a certain CDISC document.

“Process Redesign”: By introducing the standards and our new EDC system, processes from variable definition to system development have substantially changed. Particularly, we have stopped preparing CRF before system development. Namely, variables deployed on the Excel sheets are imported to the EDC system, whereby data entry interface is automatically generated. Thereafter, CRFs are printed out from the system, only by one click.

“Workflow Integration”: We have prepared our own CDISC implementation guide, and the work is done in accordance with the guide. Also, we define EDC setting specification in Excel files, which are reused when setting new studies, thus forming an integrated cycle of workflow. Additionally, we are defining basic sets of variables for certain domains such as AE or DM, to be used as common modules for setting EDC, which would help catalyze the workflow cycle.
“Standard-inspired Innovation”: Inspired by the standards, we designed the EDC system, so that the data entry interface is automatically generated from the Excel files, although the idea per se is not completely new. Consequently, time for system preparation has become much shorter. Also, less work is necessary to modify the data entry interface, for all modifications are completed on Excel sheets.

“Speed and Resource Savings”: By the reuse of EDC setting specification files and automatic generation of data entry interface, time from data definition to system release is much shorter than before, particularly for studies similar to the prior ones. Also, by adopting such procedures, much less CSV is needed, expecting a significant increase of efficiency. Although these improvements are not the direct effects of CDISC standards, they were only possible because the standards were there. However, data managers and system engineers are currently devoting substantial time and effort for the processes, particularly for studies in new disease areas, waiting for further evaluation on the speed and resource savings.

“Quality Improvement” and “Strength through Collaboration”: These issues are to be evaluated, because our experiences are still limited.

Taken together, we have started to elicit the power of CDISC standards and will continue to expand the power, in order to streamline clinical study operation and clinical studies per se, under the commitments to promote clinical science and medical innovation.

Acknowledgement

All achievements introduced here are the products of CDISC project team at TRI center, which was initiated by Director Masanori Fukushima. Also, I express special thanks to Ms. Takako Jono, who is the chief information officer at our institute, for the facilitation of all our CDISC related activities, and for invaluable comments to prepare this article.