Data Standardization and Warehousing – Two Case Studies

The testing phase of Clinical Development takes years to complete and can result in clinical trial data being collected from up to 30 or 40 clinical trials. The preparation of a submission to Regulatory Agencies requires the aggregation of these clinical data into a consolidated and standardized structure.

A data warehouse is a data repository that allows for data from multiple data sources: multiple CROs, clinical laboratories, and other vendors, to be housed and subsequently standardized. Data, regardless of file type, is imported and converted into a common format (CDISC SDTM, ADaM, or sponsor defined format). A first step in this process is the development of a standardization map, which documents the relationship of the original data to the new standardized format and is based on the electronic annotated CRF for each individual clinical trial. Following the completion of the standardization map, the data are subjected to programmatic standardization. Once standardized, the data are available in a submission-ready format as well as being available for aggregate reporting. Data from multiple studies may be reviewed and queried which is invaluable when reviewing data for a given compound.

Case Study 1 – Determination "Late" In Clinical Development that Standardization is Required

In this experience, the Sponsor began to prepare their submission and determined that the data that they were in possession of were not in a format that could be used for the presentation of the data to the FDA. The eleven studies in their development program were conducted by multiple CROs and subjected to multiple database structures. The data the sponsor had received from the CROs included data exports in excel (that only had partial data loaded), SAS, and Access. Not only did the Sponsor have seemingly incongruent data, they also had a corporate initiative requiring the regulatory submission by the 4th Quarter of the calendar year.

PharPoint was contacted by the Sponsor during the first part of Q2 to plan and implement the standardization of the databases followed by ISE and ISS analyses.. PharPoint's team reviewed the available materials, created the standardized datasets in SDTM format, analysis outputs, and Item 11 documentation (CRTs) as required by the FDA. The result of these efforts was a successful submission and attaining the Sponsor's corporate objective.

Case Study 2 – Ongoing Implementation of Data Warehouse Initiated Early in Development

In this case study, the Sponsor completed their phase 1-2 development program, which consisted of 25 studies, using multiple CROs and subsequently multiple database structures. The data sources available from these trials included, exports in excel and SAS files which were created from DataFax, OC, and Clintrial databases.

PharPoint discussed warehouse solutions with this client during Phase 2 and data warehousing and standardization activities were initiated. PharPoint converted existing databases into a standardized structure and new databases were set-up using the newly created standard structure.

The development program timeline requires analysis of the Phase 3 studies, ISS and ISE analyses, submission ready Item 11 documentation to be completed, reviewed, finalized, and incorporated into corresponding reports , within two months of database lock. This will allow the sponsor to initiate full submission to the Agency within three months of Phase 3 database lock.

PharPoint research has performed data standardization and data warehousing for multiple clients. To discuss how our experienced team may be able to help you with your clinical programs, please contact us at 919-929-0012 or by email at bizdev@pharpoint.com.