

eSource to SDTM: Trade-offs and Pay-offs

Swapna Pothula, SGS Life Sciences, Mechelen, Belgium
Donald Benoot, SGS Life Sciences, Mechelen, Belgium
Annelies Van Zeveren, SGS Life Science, Mechelen, Belgium

ABSTRACT

The advent of digital data, global networks and trustworthy cryptography have made data universally available. There is no longer a need to create physical copies of data for different parties. With a proper eSource solution in place, electronically captured data are shared real-time from the capture device to the eSource database and transferred subsequently into the clinical SDTM database, dissolving the boundary between source and CRF.

When looking at eSource systems from the traditional paradigm of paper/(e)CRF clinical trials, one might easily fall into the trap of seeing the eSource system simply as a digital replacement for paper source. However, eSource systems must be seen as part of a larger entity that integrates data capture, site automation, sample management, data flow facilitation, data visualization and data management.

For an eSource system to be integrated into the workflow requires a thorough knowledge of the technical specifications and system possibilities. An integrated eSource system causes different departments to work differently and more closely together, as it changes the inter-departmental functional roles. Streamlining the internal company structure facilitates the setup of an eSource aimed at producing quality SDTM. The implementation of an eSource system will have a profound impact on the existing workflow, where we have to decide on important trade-offs to reap the resulting pay-offs.

While we are all trying to achieve the same end result, which is creating high quality SDTM compliant databases, how we achieve this given the parameters and constraints is of equal importance if not of more, which brings our focus back to optimization and efficiency of processes at operational level.

The case study discussed in this paper is a combined clinical trial where a single clinical SDTM database is created from two study parts with different (e) Source systems for a single study: a Single Ascending Dose (SAD) part using an eSource system and Multiple Ascending Dose (MAD) part using an eCRF system. The challenges and three possible scenarios for creating a single merged SDTM database, inclusion of external data into the SDTM database and creation of submission package in the end are discussed.

The successful implementation of this case study is majorly due to SGS SDTM workflow that was designed and built to allow scalability based on requirements of clients, the in-house tools that have optimized each of the data processing step to create SDTM database and our experience with (e)Source systems.

To conclude, we had to decide on some trade-offs in task assignment, task execution, trial management and cooperation between different departments. On the other hand, increased speed, efficiency, database consistency and integrate the eSource workflow in our existing SDTM-based workflow, were some of the direct pay-offs.

INTRODUCTION

The data generated during clinical trials comes in a wide range of formats and may be generated from a variety of sources. In recent years, clinical trials have become increasingly complex involving combined study designs, a larger number of sites across a wider range of countries, as well as a greater variety of investigational technologies. Collecting, verifying, cleaning and merging this data into the desired data format, which is accurate and has the continuity to submit to regulatory bodies for approval, requires deep understanding and expertise. This means that the data capture, transfer and handling needs to be extremely flexible, while also taking into account the need to accelerate the time to market and observe budgetary restrictions.

Traditionally, source data for clinical trials was collected on paper, after which the data was transcribed in a paper Case Report Form (CRF). To ensure the quality of the CRF data, a monitor performed Source Document Verification (SDV) on the transcribed data. Only then the paper CRF could be physically transferred to the Data Management (DM) department. The Data Manager would either manually review the data by investigating the paper CRFs or the

PhUSE EU Connect 2018

CRF data would be entered in the Clinical Data Management System (CDMS). When entering the data in a CDMS, the data quality of the clinical trial was guaranteed by entering the CRF data in duplicate into the CDMS to avoid the high risk of transcription errors. Although the latter is a robust workflow, in the end it is labor intensive, inefficient and leaves room for transcription errors.

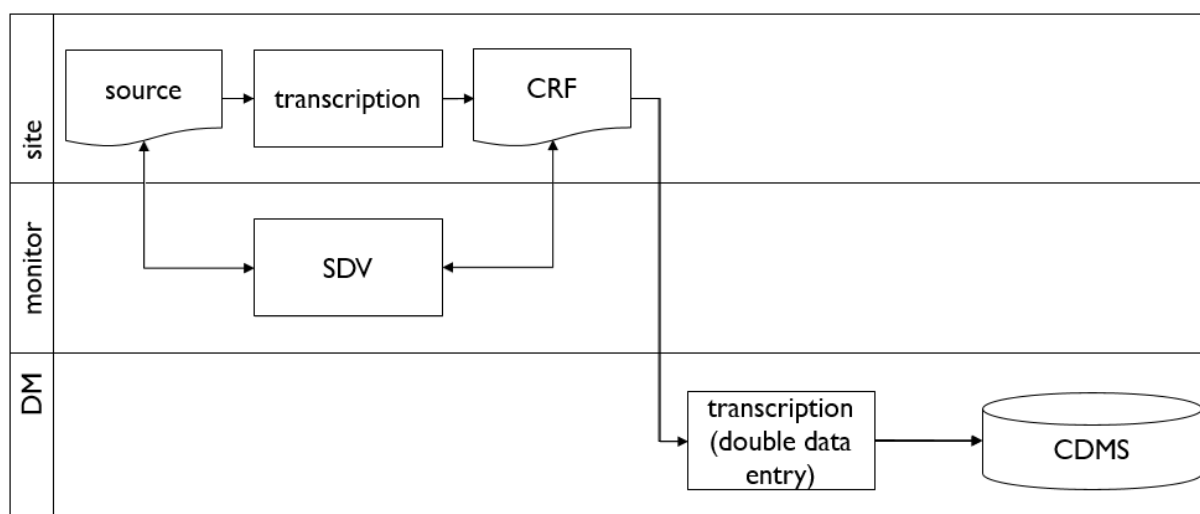


Figure 1 - The traditional paper source and paper CRF flow. Although it is a robust data flow quality-wise it lacks efficiency by having lots of time-consuming steps, such as transcription, SDV, double data entry...

With the advent of digital data, global networks and trustworthy cryptography data has become universally available. There is no longer a need to create physical copies of the data for different parties. This opened the door for the development of systems to increase the efficiency and quality of data capturing. Instead of transcribing the paper source data into a paper CRF, an electronic CRF system was used in clinical studies. Although the time-consuming transcription and SDV steps are still required on site, DM can now start processing the clinical data much sooner on the extracted eCRF data, due to the digital transport of data.

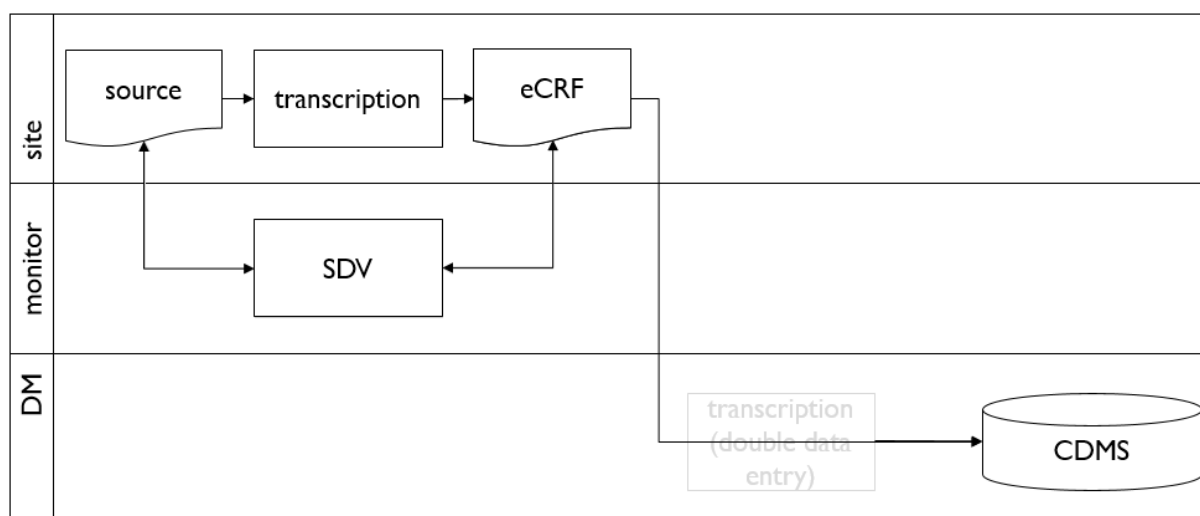


Figure 2- Evolution to an electronic CRF. The data is entered in an electronic CRF and transferred digitally to DM, increasing the speed and the quality. However on site, the process is largely unchanged and still needs time-consuming transcription and SDV.

The logical next evolution is digitizing the data flow on site by implementing an electronic source (eSource) flow. When looking at eSource systems from the traditional paradigm of paper/eCRF clinical trials, one might easily fall into the trap of seeing the eSource system simply as a digital replacement for paper source. Although an eSource system definitely performs this function, it must be seen as part of a larger entity that integrates data capture, site automation, sample management, data flow facilitation, data visualization and data management.

With a proper eSource solution in place, electronically captured data are shared real-time from the capture device to the eSource database and subsequently transferred into the clinical SDTM database, bypassing the need for any

PhUSE EU Connect 2018

transcription and SDV. In a sense, the source becomes the CRF, forcing us to rethink how we will comply with the regulatory documentation requirements.

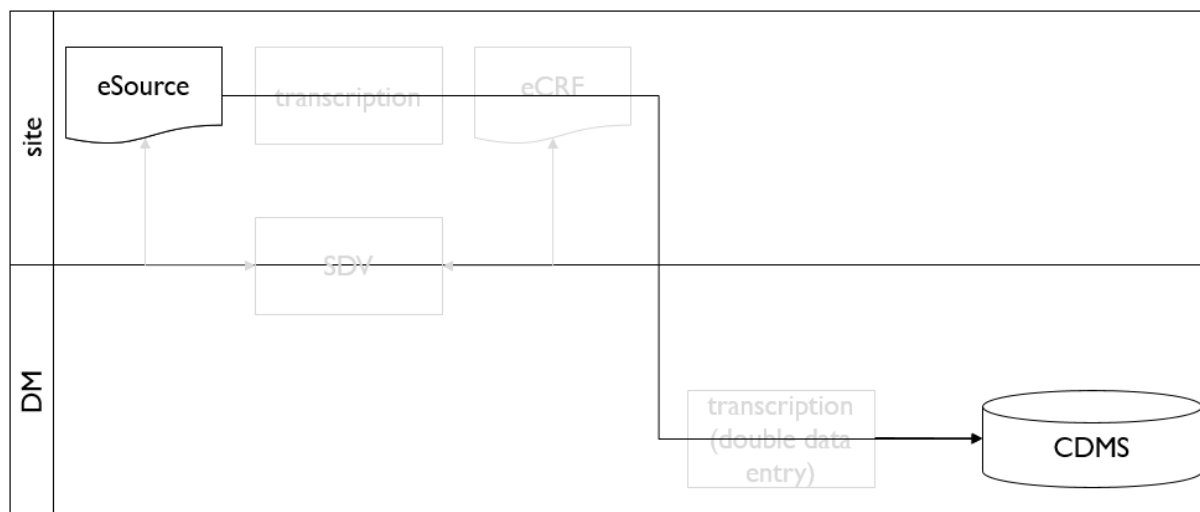


Figure 3 - Implementation of eSource. All data capture is performed digitally, so no time-consuming transcription, SDV and double data entry are necessary anymore. Additionally, the site automation further enhances the data quality.

The implementation of an eSource system, and the fundamental changes in data acquisition that go along with it, will have a profound impact on the existing workflow of the Clinical Pharmacology Unit (CPU) and the DM department. This forces us to decide on important trade-offs to reap the resulting pay-offs.

STRUCTURAL IMPACT: INTER-DEPARTMENTAL COOPERATION & TASK EXECUTION

An integrated eSource system causes different departments to work differently and more closely together, as it changes the inter-departmental functional roles. Tasks which were traditionally purely one department's responsibility (e.g. the review of the source setup by the site), now can be an inter-departmental task (e.g. DM reviewing part of the eSource to make sure the values used are SDTM compliant). Therefore, it is important to streamline the internal company structure.

The implementation of an eSource system also influences the execution of several tasks in the organization. Performing these tasks in a sensible way, in accordance with the eSource workflow and SDTM compliance, will enhance efficiency and avoid unnecessary or duplicate work. For example, tasks can become redundant because of the implementation of an eSource system (e.g. SDV, CRF design,...). The freed-up resources (often freed-up from menial tasks) can be redirected to more productive tasks.

TECHNICAL IMPACT: EMBEDDING ESOURCE IN THE WORKFLOW

Knowing the technical specifications and system possibilities allows one to optimally integrate the eSource system into the workflow. At the SGS Clinical Pharmacology Unit (SGS CPU), the eSource system includes real-time electronic data capture, site automation, sample management and direct automated data transfer to the DM department.

In eSource systems, most of the data captured electronically, and only a very limited transcription (of questionnaires or diaries) is necessary making SDV almost negligible. This allows the monitors to focus more on-site quality and source review. Additionally the esource system supports the query management of the monitor and site staff.

The eSource system automates the planning of the trial, manages the sample acquisition and tracking, and guides the site personnel through the trial assessments, increasing the initial quality of the data.

Having direct data transfers from the eSource system to DM, virtually bypasses the CRF, making the process fast, efficient, and error-proof. Because the source in a sense becomes the CRF, certain SDTM specifications should be incorporated already at site level to ensure compliant SDTM data can be created. For example, the eSource should be created using correct SDTM compliant codes and text values. These specifications ensure a fast conversion of the eSource data to SDTM format, which in turn allows us to reuse our internal data management tools, which are based on the SDTM format.

PhUSE EU Connect 2018

TWO (e)SOURCE SYSTEMS, ONE SDTM: CASE STUDY INTRODUCTION

Mapping single eSource extracts to SDTM compliant databases can be a challenge at times. We at SGS were further challenged by one of our clients, requesting creation of a single SDTM database for a study set up using two different source systems, with a variety of external data coming from multiple providers.

This study had two parts Single Ascending Dose (SAD) using an eSource system and Multiple Ascending Dose (MAD) using a EDC system due to the subject population and budget constraints, leading to two different CRFs, two different (e)Source extracts and multiple external data providers for a single study. All these added new dimensions and challenges to the process of SDTM database creation, inclusion of external data into the SDTM database and creation of submission package in the end.

The DM group at SGS is equipped with a set of automated tools and optimized workflows that make the whole process of creating SDTM databases simple and efficient. The focus of the case study will lie on the proactive approach that was devised with the end goal of creating a single merged SDTM database given the differences in the setup of the SAD and MAD using the existing processes and in-house tools.

To facilitate the explanation of the case study, the SGS SDTM Plug and Play work flow and it's adaptation for eSource trials will be explained first.

THE SGS SDTM WORKFLOW

The SGS SDTM workflow is designed and built around the plug and play concept to allow customization and to accommodate a variety of EDC / (e)Source systems, client standards and external data sources. The fig.4 depicts the SGS SDTM workflow concept where data from a variety of sources is collected and processed leading to an SDTM compliant database at the end. Based on the type and the source of the data, the way the data is processed varies.

For example as a standard, the local lab safely data coming from local laboratories is always uploaded into the lab database using the dataloader tool and if needed is coded using the coding tool. This standard procedure creates lab data that has been processed in the required way before the SDTM mapping is done.

The data coming from central labs specially blinded data follows a different path given the nature of the data. Blinded data that are provided by the labs are processed at SGS with the secure data office concept. In this case, the data coming from the lab are provided to Secure Data Office (SD Office). This is independent dedicated group within SGS that handles all data that can potentially unblind the study. SD Office works in a secure environment with separate database and server folders with limited access. Other employees within SGS cannot access data handled by SD Office.

SD Office provides SGS DM with blinded data which is loaded into our database using the data loader tool. SD Office empties all the variables that could possibly cause unblinding before providing SGS DM with the identifier transfers. These identifier data are not yet part of the SDTM database, but are used to check the structure for SDTM compliance and the content for inconsistencies well in advance while still remaining blinded.

The idea is to have a different processes process the data coming from different sources. The processing of data prepares the data to be converted or mapped to SDTM using the electronic data conversion as indicated in the Figure 4. Since these processes are all independent of each other the data are processed and mapped to SDTM as and when they are available, enabling us to generate SDTM database right from the setup of the source systems. This means the setup of metadata, generation of Define.xml and data cleaning activities are all performed on the SDTM database.

In practice we map the TEST data to SDTM. It is this concept of plug and play of processes that has enabled us to have SDTM database set up in 3 different environments: SETUP, QC and LIVE. These independent processes when plugged in together based on the requirements help us leverage both flexibility and efficiency.

PhUSE EU Connect 2018

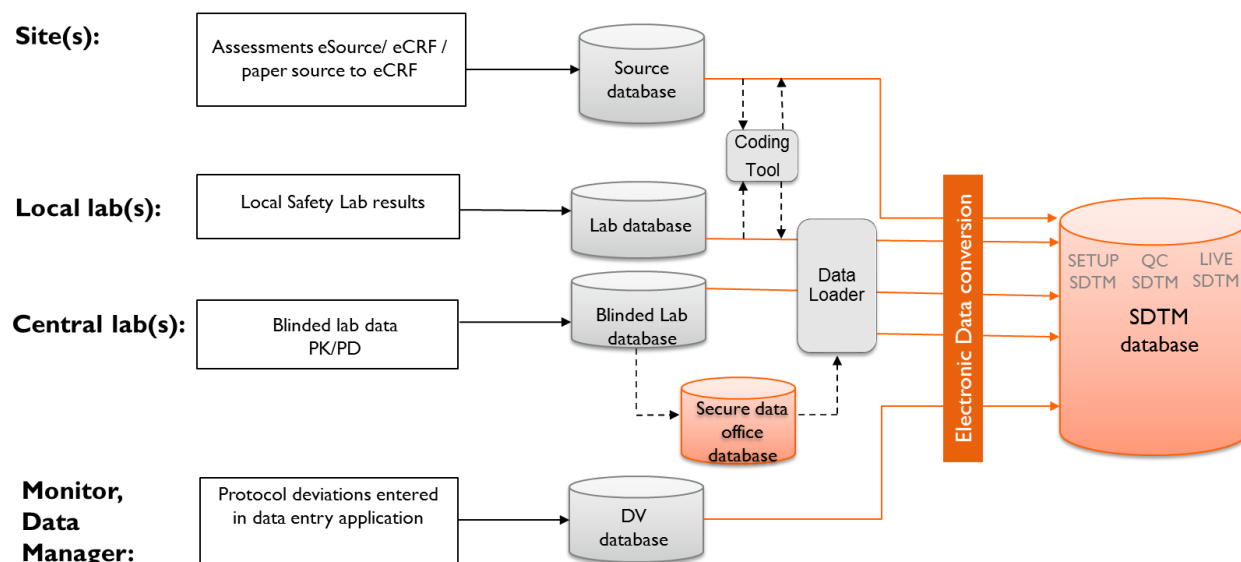


Figure 4 - SGS SDTM Workflow

THE SGS SDTM WORKFLOW WITH (e)SOURCE SYSTEMS

When SGS had to use the (e)Source systems to capture the clinical data, the SGS SDTM Workflow was tweaked a little to enable processing of data captured from a new generation of data capture system. As discussed in the beginning of the paper, we had to bring in both structural changes and technical upgrades to the existing processes and workflows to be able process and map the data captured by (e)Source systems and create SDTM compliant databases.

Figure 5 is indicative of the technical upgrades that were made to accommodate the (e)Source system. Since the (e)Source system in this case is capable of capturing local lab safety data, the lab data coming from the (e)Source system had to be processed differently. This was the only upgrade made to accommodate and integrate the capture of clinical data using the (e)Source system into our existing SDTM workflow.

The lab data and source data are now pre-processed together in contrast to the workflow as indicated in Figure 4. The rest of the processes that are part of the workflow remain the same and the electronic data conversion runs just like it does in the SGS SDTM work flow as depicted in Figure 4.

PhUSE EU Connect 2018

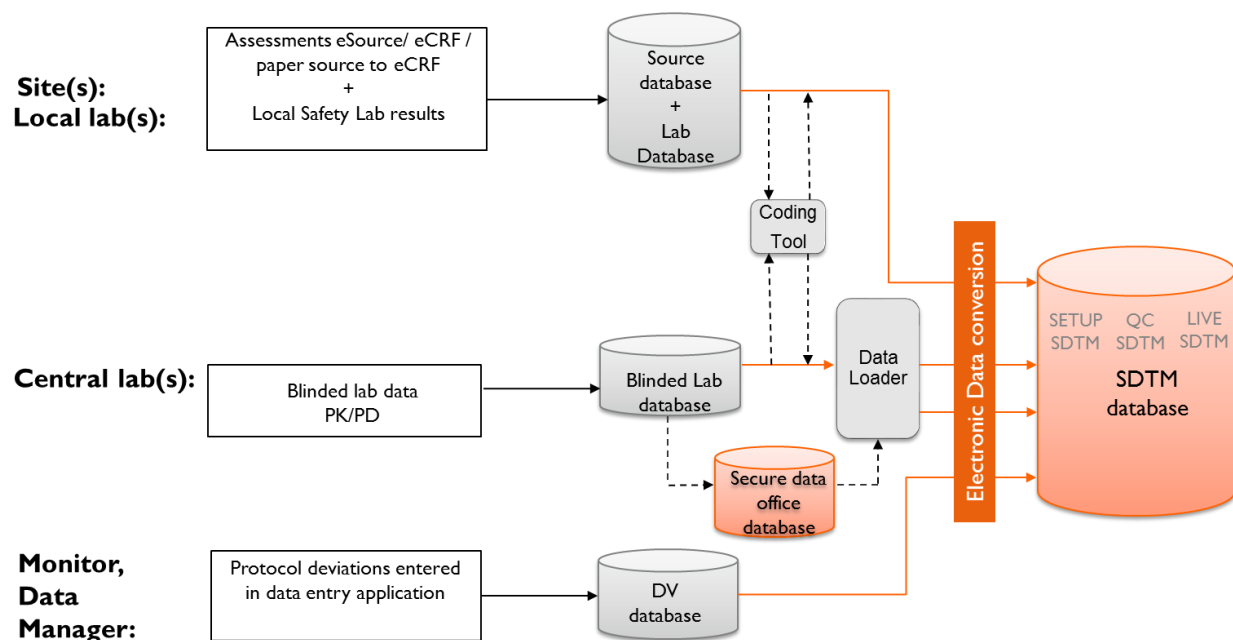


Figure 5 - SGS SDTM Workflow with (e)Source Systems

TWO (e)SOURCE SYSTEMS, ONE SDTM: A CASE STUDY

The study was a first in human, early phase trial with two parts: a Single Ascending Dose (SAD) part and a Multiple Ascending Dose (MAD) part. The SAD was to be performed on healthy volunteers at a single site. The MAD part was to be performed on subjects at multiple sites. The PK and Safety results of the SAD part were to be used to decide the doses for the MAD part. Meaning, there was a possibility that the MAD part would not be performed.

As indicated in Figure 5, integrating data capture from (e)Source systems in the workflow has enabled SGS to work with (e)Source systems with ease and efficiency. Given the benefits of (e)Source systems in comparison to the traditional data capture systems, using (e)Source systems was proving to be time and cost efficient to our clients.

Given the varied requirements for each of the parts of the study, client chose to conduct the SAD part in the (e)Source system at SGS CPU in Belgium and a traditional EDC system at multiple sites to conduct the MAD part of the study.

The challenge for the SGS DM was to set up two parts of a single study using two different data capture systems and generate one complete and SDTM compliant database. The study was to go for submission and therefore the need for define.xml.

Given the requirements and expectations of the client, the technical ability and flexibility of the workflow and time and cost constraints, SGS DM came up with the three scenarios or rather technical solutions to be able to use two different source systems for a single study. These scenarios were simulations of the implementation paths for all the possible data collection systems that could have been used for the MAD part.

SCENARIO A

The first scenario as depicted in Figure 6, is one of the possible ways in which this particular study could have been implemented. The SAD would be setup and mapped to SAD SDTM database using the SGS (e)Source workflow. The MAD data would be entered into a data entry application developed at SGS instead of using an EDC system for data entry. The data was then to be mapped to MAD SDTM database using the SGS workflow.

PhUSE EU Connect 2018

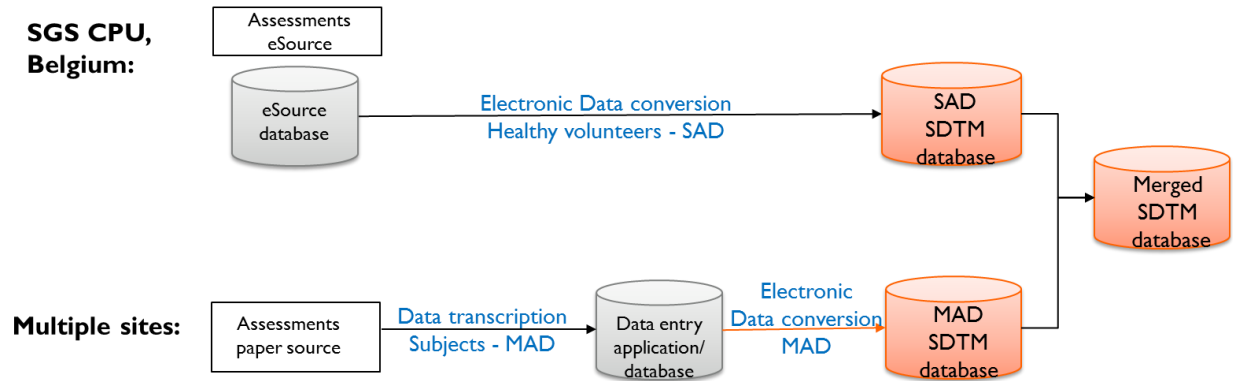


Figure 6 - Scenario A

Later on, the SAD SDTM and MAD SDTM would be merged to create a single complete SDTM database. This wasn't an ideal solution mainly because using data entry application to enter data from paper source is outdated and at the same time data entry is error prone costing both time and money.

SCENARIO B

This second scenario as depicted in the Figure 7, is another way the study could be set up, where data was to be collected from two different sources, in this case (e)Source and paper source and transcribed to an EDC system and then mapped to SDTM database.

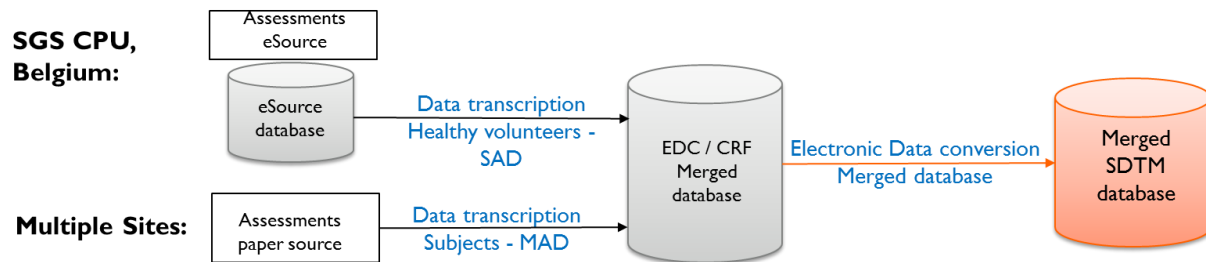


Figure 7 - Scenario B

In this approach the data was to be merged before being mapped or converted to SDTM. Although this approach would have data collected from both the parts in one place, making it easier to do the conversion and mapping of data to SDTM, it wasn't chosen. The major drawback being the transcription and hosting costs of an EDC system right from the start of the study. Apart from that transcribing data collected from (e)Source systems would not just be redundant but also result in missing out on the whole purpose and a major advantage of using (e)Source systems, as discussed in the beginning of the paper.

SCENARIO C

The scenario C as indicated in the Figure 8, was the chosen implementation model for this particular trial. This implementation model would result in setting up two different processes to process data coming in from two different source systems, in this case the (e)Source system and the EDC system.

PhUSE EU Connect 2018

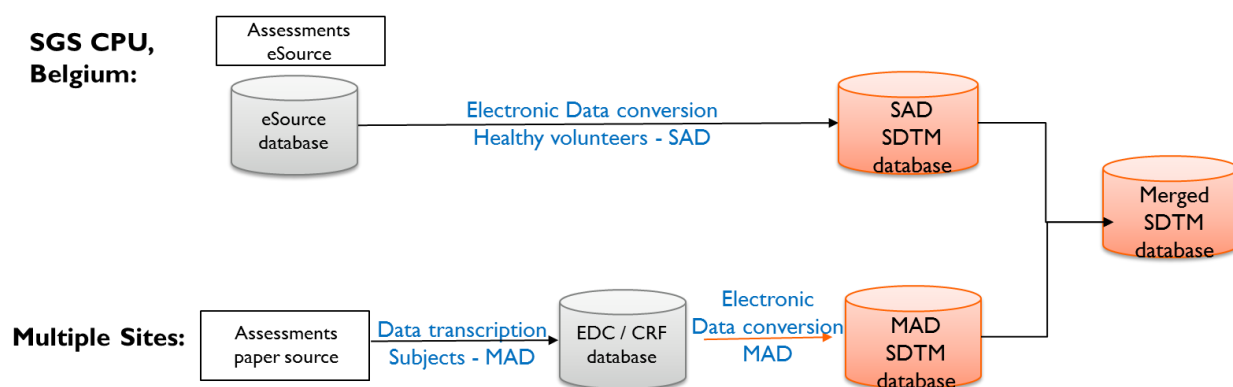


Figure 8 - Scenario C

The data would follow the SGS SDTM workflow when the source system was EDC and the (e)Source SDTM workflow when the source system was (e)Source resulting in creation of two SDTM databases one for each part of the study. The databases were to be merged later on completion of the MAD part while creating the define.xml.

This implementation model ticked all the boxes for the client with respect to cost and timelines. Also this model would make the SDTM compliant database available for analysis after completion of SAD to be able decide the right amount of dosing for the MAD part. The implementation specifics of the model in Figure 8, will be discussed further in detail in the next section.

TWO (e)SOURCE SYSTEMS, ONE SDTM: IMPLEMENTATION

The implementation of the case study at the operational level is depicted in the Figure 9, showing the flow of data from different source for each of the parts of the study. The SAD SDTM workflow is indicated in blue and the MAD SDTM workflow in green in the Figure 9. Each of the workflows have been elaborated to indicate the processes data goes through given the source, be it the source system, labs or SD Office.

The data management activities such as QC of the set up and entry of protocol deviations were performed on the SAD and MAD SDTM databases separately making these activities independent. The trial design tables were created for SAD and MAD databases separately, to be merged later on with the exception of TS table. A single TS table was made and updated as and when needed based on how the study conduct progressed.

This way all the tools that were part of the SGS SDTM workflow could be used on the two parts without major need for change in the tool design at the conceptual level. Few minor changes were however inevitable at the contextual level to be able to accommodate the requirements of the specific trial.

The two CRFs generated from the two source systems were different. The two CRFs were annotated independent of each other keeping in mind to be consistent where possible. The idea was to have the SAD setup independent of the MAD so as to be able to deliver SDTM on demand, at same time foreseeing the case when client choose not to go ahead with the MAD part (based on the analysis results of the SAD). In which case we had to have the SAD database ready to be locked with the Define.xml in place.

PhUSE EU Connect 2018

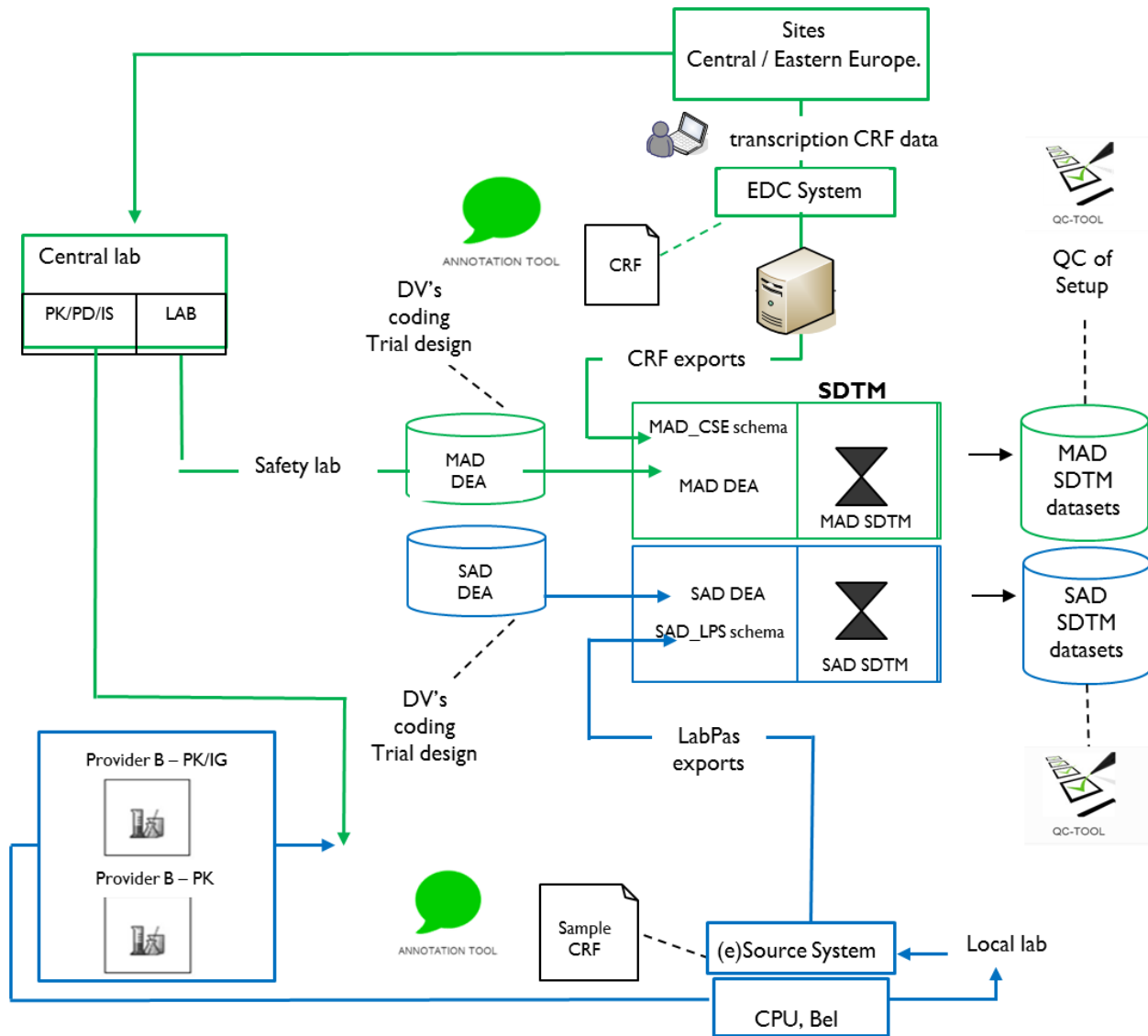


Figure 9 - Setup of the SAD and MAD workflows

The SAD and MAD databases were merged to create a merged database which was used to for data cleaning and reconciliation. Given the design and function of the in-house tools at SGS, it was ideal to run the cleaning and validation tools on the merged database. This workflow setup enabled data to flow through the standard processes as and when available and generate SDTM datasets on demand. The same approach was the case with both blinded and unblinded data from the providers.

PhUSE EU Connect 2018

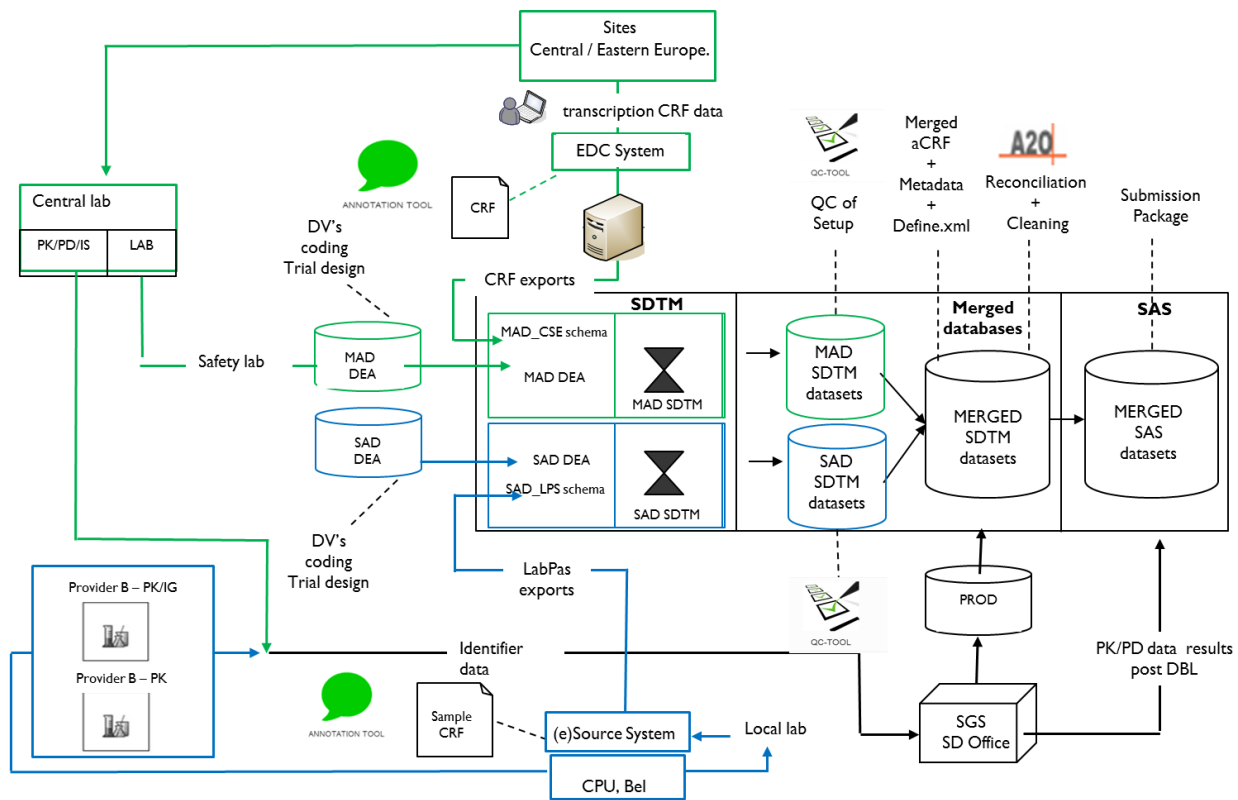


Figure 10 - Set of SAD and MAD Workflows to Create Merged Database

As indicated in the Figure 10, trial specific metadata was set up on the merged database to be able to generate the Define.xml and SDTM on demand in this case right from the setup on test data and on weekly basis.

Both SAD and MAD part of the study were conducted as planned. The database lock was performed on all the three databases, the SAD, MAD and MERGED. All the post lock activities such as release of PC, PD (Pharmacodynamics), and PP datasets were performed on the merged database. The submission package was also created on the merged database. The CRFs from SAD and MAD part were merged to create a single aCRF for submission. All the trial specific metadata updates were performed on the merged database. In the end were able to deliver a complete merged database that was SDTM compliant and submission ready.

It is the plug and play concept behind the SGS workflow and the in-house tools that enabled us to take this challenge and complete it successfully. It is the flexibility of the SGS SDTM workflow model than enabled us to implement a complex study with data captured from two different source systems in a very cost and time efficient manner, a win-win situation for both SGS data management and the client.

This definitely is one of the ways our experience with (e)Source systems implementation paid off. Since we had already accumulated enough experience with the (e)Source systems we were able to scale and upgrade the SDTM workflow at SGS to meet client request.

CONCLUSION

During the implementation of our eSource system, we determined the structural and functional impact on our workflow. We had to decide on some trade-offs in task assignment, task execution, trial management and cooperation between different departments.

Increased speed, efficiency, and database consistency were some of the direct pay-offs. We were also able to integrate the eSource workflow in our existing SDTM-based workflow, which was another pay-off. In addition to this we are also able to process multi-source trials as discussed in the case study. The lessons learnt from the setup, database lock and submission contributed to further process improvements.

PhUSE EU Connect 2018

The successful implementation of this case study is majorly due to SGS SDTM workflow that was designed and built to allow scalability based on requirements of clients, the in-house tools that have optimized each of the data processing step to create SDTM database and our experience with (e)Source systems.

For SGS, this isn't the end of the road. We have reached the technical limits of the eSource system we are currently using and, armed with the above mentioned knowledge and experience, we are in the process of moving to a next generation eSource system.

CONTACT INFORMATION

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

Annelies Van Zeveren
SGS – Agriculture, Food and Life
Clinical Research
Generaal De Wittelaan 19A
B- 2800 Mechelen
Email: annelies.vanzeveren@sgs.com
Web: www.sgs.com/cro

Swapna Pothula
SGS – Agriculture, Food and Life
Clinical Research
Generaal de Wittelaan 19A Bus 5
B-2800 - Mechelen
Email: Swapna.Pothula@sgs.com
Web: www.sgs.com/cro

Donald Benoot
SGS – Agriculture, Food and Life
Clinical Research
Generaal de Wittelaan 19A Bus 5
B-2800 - Mechelen
Email: Donald.Benoot@sgs.com
Web: www.sgs.com/cro

Brand and product names are trademarks of their respective companies.