PhUSE 2014

Paper IS01

Clinical Study Data Processes @ LEO Pharma

Jørgen Mangor Iversen, LEO Pharma A/S, Ballerup, Denmark

ABSTRACT

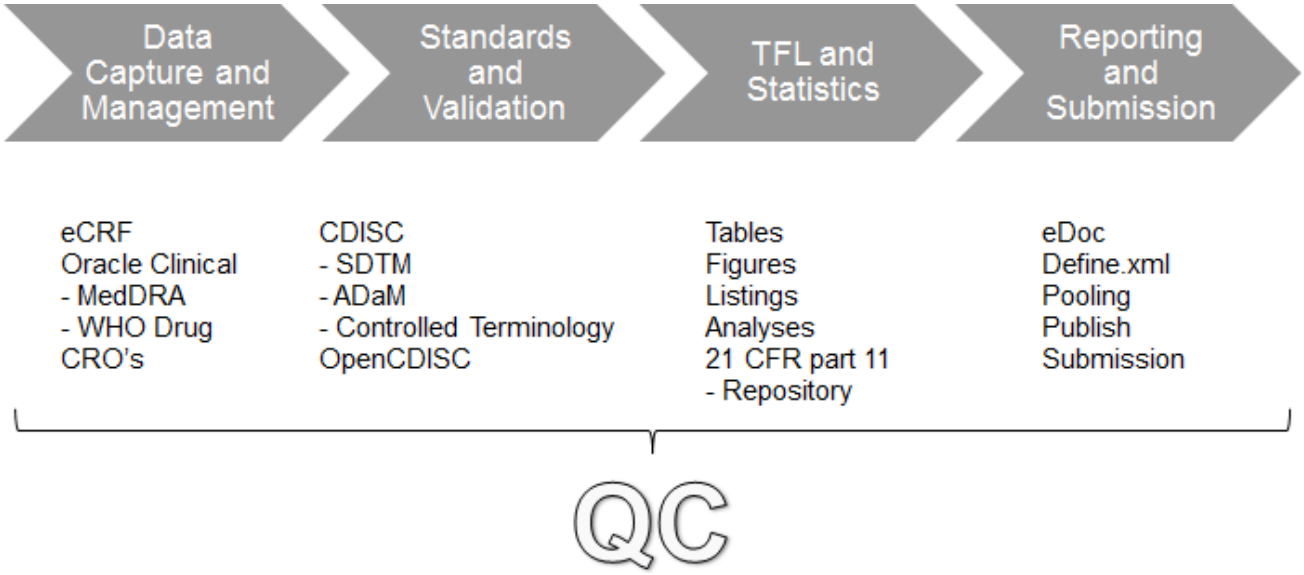
Clinical data flows through a surprisingly complex set of transformations and processes from the point where it is collected in a Case Report Form and until it is presented in a Clinical Study Report. At LEO Pharma, several different data standards from the CDISC organisation and others are used underway, some supported by standard tools, others less so. Data is validated, transformed, reported and QC'd by both standard tools and systems as well as many thousands lines of custom code, none of which may in any way alter the contents of the data, but all of which must act to ensure that the same data remains correct and well documented, while being used for many diverse purposes. Covered here are details of which data standard is used and for what purpose, which standard system controls the flow of data, and the amount of custom code.

INTRODUCTION

LEO Pharma made CDISC standards part of its internal processes in 2004, documented in a response to the FDA where LEO got approval of submitting data electronically in SDTM format of the time, making LEO one of the first four companies to do so. All new studies since 2004 have benefitted from the CDISC data standards, which have been used for all work on clinical standards. Some systems in our processes are using other data standards internally, but only strictly interpreted CDISC standard data travel between systems.

PROCESS OVERVIEW

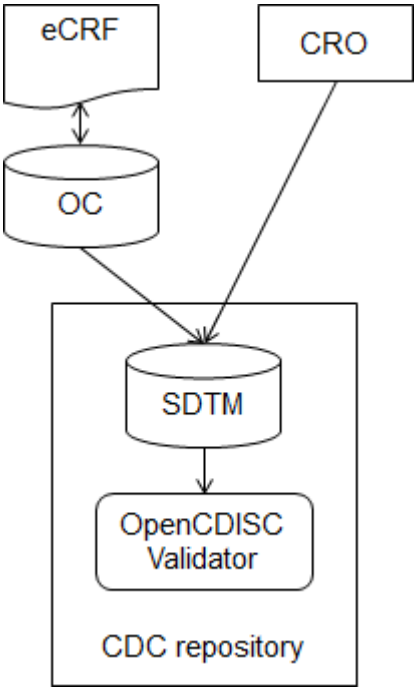
Data travels through several departments and organisations and through several IT systems as well. The emphasis will be on the systems and methods, as organisations tend to fluctuate more than systems.



This diagram depicts the main processes that handle the clinical data, detailing briefly the content of each process. There is a host of other processes to define and set up a clinical study, but these are the ones dealing explicitly with clinical study data.

DATA CAPTURE AND MANAGEMENT

Data capture uses a two-tier strategy. We can choose to set up the study in Oracle Clinical (OC) and use the eCRF system of Oracle to collect data. CRF pages are built in OC, and data management subsequently takes place in OC as well, including coding of terms to MedDRA and WHO drug coding systems. Finally the data is exported from OC as SAS datasets in SDTM format. This is not a native functionality of OC, so some programming is needed to accomplish this.



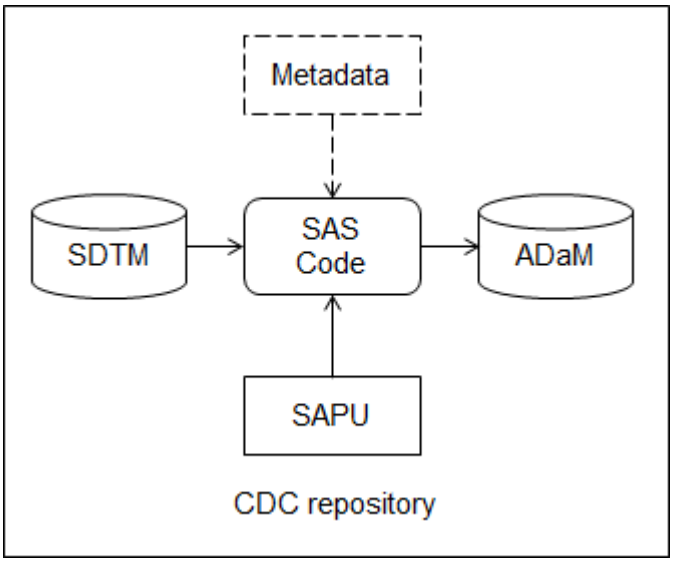
Alternatively, we can choose to outsource data capture and data management to a CRO, specifying SDTM as the data delivery. The CRO is then expected to perform exactly the same tasks, delivering the same quality of data, but using their own procedures.

In either case, the SDTM standard is a crucial element in securing that data from either source are of sufficient high quality. Leo's primary tool to validate the quality of SDTM data is the OpenCDISC Validator. This tool requires data to be delivered in- or converted to- SAS transport files. This is the first place where we reap the benefits of insisting on a high level of adherence to standards. When data always exist in a uniform data standard, we can use the same methods, tools and systems in subsequent processes.

As indicated, when the clinical study data reaches the SDTM format, it is imported into a clinical data repository, a controlled GCP validated environment, within which all subsequent processes takes place.

STANDARDS AND VALIDATION

At LEO we have chosen that the main location for complex data transformations is in the creation of analysis datasets (ADaM). In our interpretation, SDTM is much focused on the collection of data, and ADaM is equally focussed on the reporting and analysis of data. One key factor in reducing this complexity is to ensure that both the SDTM format and the ADaM format of essentially the same data vary as little as possible from the standards defining them. This way the transformations of data can be reused from study to study with minimum change of the code.

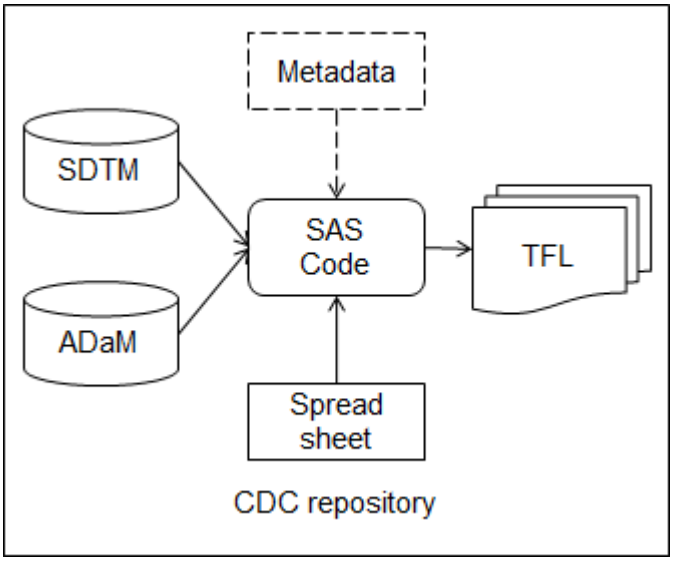


As the standards primarily deal with the format and structure of the data, and not the contents of the data, the standards are flexible enough to allow for very different studies to be held in the standards.

Keeping all processes inside the repository ensures that all study data are 21 CFR part 11 compliant.

TFL AND STATISTICS

The production of tables, figures and listings (TFL) and statistical analysis is based primarily on ADaM data sets, and to a lesser degree on SDTM. Beyond the clinical study data, a small amount of metadata detailing the appearance of tables and some spread sheets governing the production of tables and listings is used for TFL.

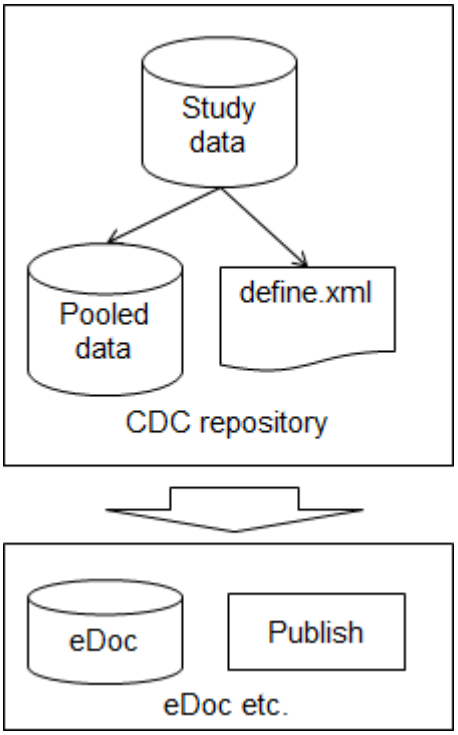


The production of tables and listings relies heavily on standard macros, making this process standardised, configurable, reusable and still flexible, as the macros are parameter based, partly automated and modifiable for special needs. This level of macros does not exist for the production of figures, although plans for such macros exist. Statistical analysis is by its very nature very study specific, and is thus performed on an ad-hoc basis as a rule.

The described level of macro use is possible only as a direct result of data being highly standardised. Also here, keeping all processes inside the repository ensures that all study data are 21 CFR part 11 compliant.

REPORTING AND SUBMISSION

While all studies are reported in a Clinical Study Report (CSR), not all studies are submitted to authorities. Some studies may have pure marketing rationales, and thus serves other purposes than to be submitted to authorities.



As very few studies get submitted by them selves, there is often a need to pool any given study with other studies. This is the point where it really becomes evident if any studies in a pooling deviates from standards. When merges and/or joins fail, it's the compelling evidence of nonconformity. In the rare cases where all (or at least most) of the studies to be pooled *do* comply with standards, a submission can be compiled in a matter of a few weeks. Most times it takes several months to compile a submission, and most often due to the need to rework clinical study data to comply with CDISC standards. This also applies when the pooling is to be submitted in a different (older or newer) version of standards than the studies involved were originally made to comply.

This observation leads to the identification of a need to roll any given clinical study data up and down the ladder of standard versions in a uniform and systematic way, in order to avoid that todays new clinical study data doesn't become tomorrow's legacy data.

The picture also shows a mechanism to transfer clinical study data and TFL's from the 21 CFR part 11 compliant repository into the equally compliant document handling repository without any possibility of tampering with the documents. Furthermore, it shows the production of define.xml as documentation of the clinical study data.

QC PROCESSES

The QC process is an integral part of all the other processes, and *not* something added at the end. Every step the clinical study data passes through requires some form of QC before the data is passed on the next step. As a general rule, LEO uses a risk based strategy for QC. This means that the most important and critical parts of the clinical study data, as well as the primary and secondary endpoints may be subject to an independent parallel programming, while other parts may rely on validated standard macros always to deliver correct results on known (standard) inputs. Important parts of a study include the CRF, SDTM, ADaM, tables, listings, figures, analyses, pooling, define.xml and everything else relevant. One particular experience that LEO has made is that any deviation from standards requires additional QC work. Greater adherence to standards yields shorter production time of input to the CSR where months could be reduced to weeks of production times.

CONCLUSIONS

After 10 years of experience with CDISC standards, many details are still inevitably hand held. However, the processes are 21 CFR part 11 compliant end-to-end, and our strong focus on standardisation yields a high level of automation, great flexibility for agile adaptations of high-speed, high-importance projects, as well as a rather quick production of the CSR components.

The clinical study data processes are a work in progress aiming at automating also the protocol, CRF, figures, some analyses, define.xml, pooling, submissions, data standards upgrades and possibly more.

CONTACT INFORMATION

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

Jørgen Mangor Iversen

LEO Pharma A/S

Industriparken 55

DK-2750 Ballerup

Denmark

+45 72 26 31 16

joergen.iversen@leo-pharma.com

http://www.leo-pharma.com/