

Bimonthly newsletter of XML4Pharma,
Schlossbergstrasse 20, DE-78224 Singen, Germany
Phone: +49 7731 975044
Web : www.XML4Pharma.com
Mail: Info@XML4Pharma.com
June-July 2010



**XML4Pharma is a CDISC
Registered Solutions Provider**



CDISC European Interchange

The CDISC Interchange in London was a great success. Although especially some US speakers didn't make it due to the volcano ash problems ten days before, the organizing committee was able to present a fantastic program.

For me, the conference started with a short German User Group meeting on Tuesday evening, followed by a reception on the highest floor of the hotel with a magnificent sight over London.

Wednesday morning had an early start, as we still had to build up the booth for the exhibition (though it was promised to us that we could already build up on Tuesday), but everything went smoothly thanks to the hotel technician, who even provided adapters for the electrical connections (speaking about standards ...).

After the keynote presentation, there was an extremely good session on CDASH. It is extraordinary how CDASH has become a real "must" in clinical data collection, and how it has so quickly be embraced by many companies, as well for paper studies as in EDC.

Wednesday also saw the ODM "breakout discussion session". Although we were a small group, it was a very good discussion, from which I can take a lot of new ideas for new features and improvements for a next version to the ODM development team.

The social event at "Madame Tussauds" was great. I finally met David Beckham and many other famous persons ... at least their wax figure.

I had my own presentation on Thursday about the upcoming ODM-extension for trial design, adding features such as arms, epochs, inclusion/exclusion,

workflows, timings etc.. I also demonstrated how such an ODM file with trial design information can be used as a source for automatically setting up a caBIG Patient Study Calendar. One important conclusion was that though both are implementations of the Protocol standard, this does not guarantee that these are 100% interoperable.

The ODM and SHARE sessions were also very interesting. For me, it was especially important to see that big sponsor companies like GSK and Lilly are basing more and more of their systems on ODM, with the strong desire to integrate ODM and SHARE. This is surely an area in we will see some things happening in the future.

The last presentation was given by Wayne Kubick (Phase Forward), CDISC's SDTM pioneer and veteran. Using his great sense of humor, he shared with us his concerns about using HL7-messages and formats for porting submission data (among others), something I am also very concerned about (see previous newsletters).

On Friday, I attended the ADaM training. It improved my understanding of the ADaM standard a lot, but it also made me realize how intensively the standard is interwoven with SAS software, or at least very strongly influenced by it. So I am more and more asking (though I have not found an answer yet), whether ADaM is really an open, vendor-neutral standard.

All together, this was a great event. We had many interesting presentations, many interesting discussions, the booth was very well visited, and we had good weather and a great social event.

What more does a CDISC Interchange need?

ODM Version 1.3.1

CDISC has just [published the ODM v.1.3.1 standard](#). This version is a minor update (it uses the same namespace as ODM 1.3) of the well-established ODM v.1.3. The XML-schema has been improved further, for example by allowing the “Alias” element on almost all “Def” elements, and also allowing zero occurrences of “Ref” elements within their parent “Def” element. The former is especially important to allow the ODM to port annotations, e.g. for SDTM, for CDASH, and for other content standards (see also [our Wiki article “Annotating ODM with SDTM and CDASH Information”](#)), and for providing links with electronic health systems and hospital information systems. The latter is especially important for giving the ODM even better support for extensions, not only “vendor” extensions, but also for the upcoming “Trial Design” extension, as well as for future extensions (SDTM? SHARE?).

Also the documentation has been further improved, with more, and especially more clear explanations and examples on “difficult” elements such as the “Include” element.

Version 1.3.1 is fully downward compatible with version 1.3.

A new, strongly improved version of the ODM Viewer is now available

Inspection of (very) large ODM files in a user-friendly way is not an easy task. When using classic technologies (such as the “Document Object Model”), viewers get difficulties with memory usage (thus slowing the viewing down). Therefore we recently started using a new technology in our ODM Viewer, which has a much lower memory footprint when inspecting large ODM files in a user-friendly way.

This technology is called “[VTD](#)” (“Virtual Token Descriptor”), and is seen as a major breakthrough in XML technology.

This has led to a new version (v.1.4) of our popular ODM Viewer. The new version is considerably faster, uses less memory, and is even more user-friendly than before.

Existing users of the ODMViewer can get a free upgrade. Just send us a short e-mail, referencing your existing license.

If you would like to “try-before-you-buy”, then you can [download an evaluation version from our website](#).

Using the new version, we could easily display and search through the (large) sample library that was published by Eli Lilly on the CDISC website ([members-only section](#)). We were even successful with much larger ODM files that we obtained from our customers.

At the same time, some more new features were added and some extra improvements were made. The real value however is the increased in speed and the lower memory usage.

CDASH User Guide for public review

CDISC also recently published the “[CDASH User Guide v.1.0](#)” for public review. Though it has almost 200 pages, it was a real pleasure to read and review it.

For me, this “User Guide” is a fantastic document, clearly explaining many of the use cases for CDASH implementation, and also showing as well paper- as EDC example forms. It also announces the publication of an ODM implementation of the CDASH standard, which I hope will come soon.

Thumbs up for the volunteers that wrote this document!

CDASH v.1.1 for public review

At the same time, the CDASH team also published the draft version of the [CDASH standard v.1.1](#). I would strongly recommend anyone interested in CDASH to download the document and to review it.

Although the official review period is already over, please send your comments to the CDASH team. Only in this way, we can make a great step forward again in the development of the CDASH standard.

CDISC Metadata Submission Guidelines v.1.0 for public review

About at the same time, also the draft “Metadata Submission Guidelines” for SDTM has been published on the [CDISC website](#).

Unlike the CDASH UG and CDASH 1.1 drafts, I am not so very happy with this draft document.

I.m.o., especially the part on define.xml is counterproductive. It explains how the (HTML) view of a define.xml should look like after applying the stylesheet, but does not at all explain how this can be realized. In many cases, it speaks about the define.xml as it were an HTML or PDF document: it speaks about tables, columns and rows, this though a define.xml file does not have any tables, columns or rows. The latter are only the result of applying a stylesheet.

No information nor guidance however is given about

how a define.xml should be generated or structured to obtain the desired result. For this, the user is left alone.

The authors of the document have clearly not understood what define.xml is about. Define.xml is a structured (XML) document with lots of information about the submission. It doesn't however describe how that information is presented. For the latter, the stylesheet is responsible. The document however does not even try to describe the structure of the define.xml, does not describe good practices to come to it, nor does it say something about how a stylesheet can be constructed.

For me, the document can be compared to a document describing how a glossy sales folder of an expensive sportscar should look like. This does of course not help those people that actually need to build the sports car, that need to assemble it, that need to design the motor for it, etc.. Translated to the clinical world, the document is only of interest for the Regulatory Affairs people that only look at the finally **displayed** result (which could as well be an HTML or PDF file). It does however not help in any way those data managers, SDTM specialist or IT people that need to assemble the define.xml file and generate the stylesheet.

So, in my opinion, the draft "Metadata Submission Guidelines" needs a lot of rework, if not a complete rewrite.

News from the FDA front

A new [blog entry of Rebecca Kush](#) on the CDISC website about the new draft FDA document "*CDER Data Standards Plan Version 1.0*" shows that there is some evolution in the way CDER (and probably also CBER) is thinking about data standards and about a suitable format for transporting SDTM data to the agency. In our previous newsletters we did already report that both these departments have announced not to accept any submissions in HL7 format in the "foreseeable future", and will stick to SAS Transport 5 in the next years.

New in the latest document is that CDER now also states that HL7 is not ready for a transition: "*The current version of the PDUFA IT plan has caused confusion and concern among both external and internal stakeholders because it implies a shift to the submission of study data formatted in accordance with CDISC-HL7 standards within the next few years. This impression is not correct. Neither the HL7 standards, nor the Agency, nor regulated industry are ready for such a transition*". Essentially meaning that the CDISC-HL7 format is not there yet at all.

However, what is already there for exchange of clinical data (who would say that submission data is not clinical data?) is the ODM standard, extended to define.xml for submission metadata.

At least 3 major companies in the SDTM scene are using ODM to store SDTM data: only in the very last step in the process, the ODM is transformed into SAS Transport 5 for final submission to the FDA.

So why does the FDA not migrate to ODM (or an extension to it) instead of to HL7-XML (which does not exist yet) for submission of clinical data? All SDTM and CDASH controlled terminology is already available as ODM Codelists, the stylesheets for inspecting the data are available, viewer software packages are available, and ODM already implements CDASH in an excellent way. Validation software for ODM is already available, and the SDTM „rules“ can easily be implemented using XML technology (such as Schematron) in a clear, transparent and unambiguous way. SDTM submissions in ODM format can probably very easily be loaded into the Janus warehouse as the ODM format is a pretty simple XML.

The FDA would benefit strongly of such a transition to SDTM submissions in ODM format, as the ODM format eliminates almost all limitations of the currently used SAS Transport 5 format, such as the 8-, 40- and 200-character limitations, the SUPPQUAL disaster (currently huge amounts of data are submitted in SUPPQUAL domains and need to be moved back to their original domain), the COVALn catastrophe (in the Comments (CO) domain) where comments are separated from their origin, etc.. The goal of SDTM has always been that reviewers do not have to be bothered with shifting data from one place to another, but currently they are plagued by moving SUPPQUAL records and Comment records back to their original domain.

Also the industry would strongly benefit: software systems currently need to use a magnitude of „tricks“ in order to format the submission data in this ancient SAS Transport 5 format, making the process of generating SDTM (and AdaM) datasets highly inefficient. When an ODM format is used however, the time needed by a user to generate SDTM datasets can be reduced by at least a factor of 2-5. Also the cost for software that is needed to generate SDTM datasets can go down by a factor of at least 10, as much more vendors will be able to offer packages that are able to produce SDTM datasets from operational data.

A last argument (but not the least one) is that CDASH data, which are currently usually already captured in ODM format, can then extremely easily

be transformed into SDTM datasets, using XSLT stylesheets. The latter can even be developed by some of the CDISC volunteers and made available to the CDISC community for free.

SDTM-ETL™ and ADaM support

We received a lot of questions regarding our SDTM-ETL™ software and its “partial” support for ADaM. We try to explain a bit more in detail here.

The ADaM standard is strongly SAS-oriented, though it is claimed to be an open standard. For example, unlike SDTM, it uses SAS-formatted dates and times. The way these formats work is however

known: a SAS date is the integer number of days since January 1, 1960, and a SAS time is the number of seconds since midnight. So, such formatting can and is implemented in SDTM-ETL™.

However, SDTM-ETL™ software is currently limited to subject-related domains and variables – it is not a statistical package. So currently, SDTM-ETL™ will be able to generate all those ADaM datasets that are subject-related, such as the ADSL dataset and any other dataset that has a structure “one record per subject per XYZ ...). Datasets in which real statistics are performed however (such as distributions) are currently outside the scope of SDTM-ETL™.



CDISC European Interchange London 2010 – Social Event.
Amanda (CDISC Dir. Media & Communications) in company of Sean Connery – who is most famous in CDISC?

Disclaimer: all information in this newsletter is to the best of our knowledge. We are not claiming any correctness nor completeness.

Copyright: you are entitled to print out this newsletter, and to pass it to other people in electronic or paper form, except for the purpose of legal action. Using the contents of this newsletter for any legal action is strictly prohibited.

For reproduction in electronic form on a public or non-public website, you do however need our permission.