# XML4Pharma is a CDISC Registered Solutions Provider

## ODM-SDM (Study Design Model) extension to be published for public review in the next days

It is so far: after a year of hard work, the ODM-SDM team is ready to publish this new proposed standard for public review. So keep an eye on the CDISC website, and when it is there, download the documents, schemas and sample files, and ... review them. And please do not forget to submit your comments in time!

SDM (Study Design Model) is an extension to the ODM standard allowing to define features of a study design that were not covered by ODM sofar, such as:
- eligibility criteria (inclusion/exclusion) and protocol summary
- structural elements: arms, epochs, cells, segments, activities
- workflows (between activities), entry- and exit criteria for segments and activities
- timings between activities

A picture of a workflow generated by our ODM Study Designer (for the famous LZZT study) is shown at the end of this newsletter.

Already different stakeholders have shown great interest in this new standard, as it allows them to:
- search for eligible subjects in databases
- set up workflows and timings for studies in hospital planning systems and EDC systems
- automatically generate SDTM "trial design"datasets from the protocol

So we expect that the publication of this new standard will again be an important step forward for CDISC.

## ODM Study Designer ready for ODM-SDM

One of our assignments in the team was to ensure that the new standard is also implementable, as we realized that a standard that is not implementable in tools is a dead standard. So in each development cycle, we tried to implement the draft standard and XML-schemas in our "ODM Study Designer" and developed wizards for adding information such as the structural design of the study (arms, epochs, cells, segments), the workflow between activities, etc..

It was a lot of work, but it now enables us to present a new version of the ODM Designer which works with the new standard, and allows users to fully implement it. For users that want to work with the new standard, this software package comes right in time!

This updated ODM Study Designer v.2011-beta (we will call it beta as long as the ODM-SDM is not "final") is will become available in the next few days. More details can be found on our website.

## XML4Pharma's SMART Challenge submission is now life

As we announced in our previous newsletter we are taking part in the "SMART" challenge of the US Government.

We can now announce that our submission "SMART Clinical Research" has been accepted and is competing with 14 other "smart apps".

This challenge is about developing "apps" (yes, just like for your smartphone), that uses a given electronic health record (EHR) system, and does something "smart" with the data from that system.

Sample applications that were made available include an "app" that calculates cardiac risk, and an "app" that checks for drug/allergy conflicts.

Our own "app", the "Clinical Research App" takes demographic, medical problems and current and past medications information from the EHR system, and uses them to first check whether a subject is eligible for the study, and then prepopulate CDASH "Demographics" and "Prior and Concomitant Medications" case report forms (eCRFs). Especially the latter is extremely useful, as some of the patients have over 100 prior medications in their EHR. Just suppose these had to be typed over from the EHR into the EDC system.
After having completed, and if necessary corrected the data in the forms, and upon submission of the form to the server, an ODM file is generated with all the data, and a PDF is generated synchronously that can be stored in the investigator's archive. Some screenshots can be found at the end of this newsletter.

Our "app" uses servlet technology, server-side XForms, and XSLT and XSLT-FO (the latter for generating the PDF).

The very special thing about "SMART" is that the EHR and any "app" runs in a common sandbox, meaning that a user can add and remove apps (just as you do on your smartphone), and that there is a seamless integration between the EHR system and the selected "app": the user just selects a patient, and then applies the "app" to the data of that patient. So it looks to the user as the "app" runs within the EHR system! I say "looks as", as any of the "apps" in reality runs on different remote servers, usually on a server of the company that created the "app". It is even possible that a single "app" uses different servers for different parts of its application.

Also very special about "SMART" is that it does not use any messages (like HL7-messages) for exchange of patient data between the EHR system and the "app". Instead it makes an API available where the data can be queried using SPARQL, which is an RDF (RDF = Resource Description Framework) query language, returning RDF triples. As such, this new approach is surely a paradigm shift, circumventing the "jungle of standards" that exists in healthcare IT. We already now see that major EHR System vendors are developing a very similar API, instead of trying to develop export systems based on HL7 or other messages.

Another interesting "app" that was submitted by one of the competitors and which is also in the domain of clinical research is the "rxClinicalTrials" app, that takes the list of medical problems of the patient, and

searches for suitable (recruiting) clinical trials on ClinicalTrials.gov. A screenshot can also be found at the end of this newsletter.

Additional information about our submission to the contest can be found on our website at: www.xml4pharma.com/Smart_Challenge/. It also has a number of screenshots of our "app" running in the SMART container, and a short demo movie. An overview of all submissions to the challenge including demo movies can be found on the "SMART Apps for Health Submissions" website.

### CDER amendments to SDTM-IG 3.1.2

CDISC and FDA-CDER recently published a set of amendments to the SDTM-IG 3.1.2. You can find an introduction to them as well as the full details on the CDISC website. At the same time FDA-CDER published a document "CDER Data Standards Common Issues", referencing the amendments, explaining some of them, and adding additional requirements.

Though the FDA document contains a lot of information about "good practices" (which experienced mappers already knew), it also contains some frightening stuff, such as the recommendation to also submit a printable "define.pdf" file. This shows that even 6 years after publication of the define.xml (CRT-DDS) standard, CDER does still not have the instruments in place to work with it.

One of the amendments says that CDER expects that "EPOCH", "ELEMENT" and "ETCD" are to be submitted for every subject-level observation. Essentially, these are not data that are capured on the CRF, but are usually derived from the visit number and the protocol, and thus essentially should not be in SDTM (SDTM data is about captured data, not about derived ones). We do however understand the needs of CDER here, making review easier by not having to do a lookup in the trial design and subject visits datasets (can't their tools do that?)

Another CDER requirement is one about splitted datasets: "*Sponsors should submit these smaller files in addition to the larger non-split standard LB domain file*".
This really does not make any sense: first of all, in **reality**, the smaller files (e.g. LB.1, LB.2 etc.) were in reality never "split" from a larger file, but were generated as different instances of the same domain. So the "larger, non-split domain file" does usually not even exist, meaning that sponsors should start merging the different files into a huge one! Secondly, this may mean that the huge "non-split domain file" can't even be read by the FDA: "*... can*

*exceed the reviewers' ability to open the file using standard-issue computers ...*". If the FDA can't read these large files, can the sponsor? If not, does this mean the sponsor is required to generate a huge LB file that cannot be opened? I.e. are they required to submit a bag of which the contents cannot be checked?

Thirdly, it was reported to us that in case the "non-split domain file" is submitted together with the "splitted" domain files, the submission will not pass through Web-SDM, as the latter tries to (re)generate the non-splitted domain file, and then throws an error that says that there are duplicate domain files. So, personally, I think CDER should drop this amendment immediately as it does not make any sense at all.

Another problem with this amendment is that it opens the door for different "dialects" of SDTM: one for CDER, one for CBER, one for ...
**Dialects of standards are standards killers.**
My opinion is that if CDER wants to keep these additional requirement upright, these should go as "conditionally expected" into the SDTM-IG, and <u>not</u> as an amendment, but in a new, updated version of the SDTM-IG, with a different version number (e.g. SDTM-IG 3.1.**3**).

For the rest, the CDER document is an extremely useful one, and I recommend anyone who is involved in SDTM submissions or generating datasets for SDTM to read it, as it really gives an insight into what CDER expects to find in an SDTM submission.

My comments to the amendments can also be found on the "Assero blog". I have also send them to the FDA itself.


**<u>New features in the upcoming release of SDTM-ETL</u>**TM

We are also working on implementation of new features for the upcoming release of SDTM-ETL™ v.1.6, such as:
- implementation of the new (still in draft) CDISC oncology domains (TU, TR, RS)
- implementation of the CDER amendments to SDTM-IG 3.1.2
- automated creation of RELREC datasets
- batch execution of datasets from existing mappings

Especially the automated creation of RELREC datasets is a major new feature: it builds on the experiences obtained with the feature to automatically generate "Comment" (CO) domain datasets. A video movie demonstrating this new feature is already available on our website.

When the ODM-SDM extension is final, we will also add features to generate and execute the mappings for the "trial design" domains (TA, TE, TV, TI, TS). This will be pretty easy, as the ODM-SDM extension has been designed so, that automated generation of these datasets should easily be possible.

**<u>CDISC Express - the next "silver bullet"?</u>**

Clinovo, a California-based clinical services, consultancy and software company, recently announced the release of "CDISC Express", a "*free application that automatically converts clinical data into CDISC-SDTM using an Excel framework*".

**Sounds promising isn't it**?
Is this the "silver bullet" we have all been looking for? So we had a look and downloaded the tool and the documentation and started working with it.

A first disappointment was that the tool requires that the source data is already in SAS format, and that you need a SAS license to be able to execute the mappings. There is no support for CDISC ODM, which is the worldwide standard for exchange of clinical data and metadata. Of course you could transform your ODM data to SAS datasets (e.g. using PROC CDISC) but why do so when there are already (relative low-cost) user-friendly tools available on the market that allow to generate mappings starting directly from ODM itself?

In CDISC Express, all mappings are defined in a single Excel file, which needs to be populated by the user. There is no wizard or drag-and-drop tool to help the user developing the mappings. I haven't also seen any features that e.g. allow to add/insert extra allowed variables (such as is the case for timing variables) Essentially, snippets of SAS scripts need to be added to cells in the Excel worksheet, requiring the user to have ... SAS programming skills. The Excel file is then transformed into a set of SAS programs which can be executed on the source SAS datasets.

CDISC Express also promises us to generate a define.xml file automatically once all datasets have been generated, but also here, a lot of information needs to be provided in the Excel worksheet first. We didn't however find any support for value-level metadata (def:ValueList), computational methods (def:ComputationMethod) or external codelists (ODM "ExternalCodelist"). Especially the lack of support for value-level metadata is giving us a lot of concerns. Other information that is error-prone must also be entered by hand into the mapping Excel file,

such as the "Length" of the SDTM variable - it is not calculated from the information in the source or result files.

So yes, it does generate a define.xml file, but this file will highly probably still needed to be edited by hand or using special user-friendly define.xml editing tool.

Conclusion: if you do already have SAS, and your clinical datasets are in SAS format, CDISC Express may be a helpful tool to automatically generate SAS programs to transform your data into CDISC SDTM. This is however also the danger of the tool: blind trust in such a tool may lead to disaster - you still need SAS programming skills and a lot of SDTM knowledge and will still need to do a lot of QC on the generated programs.

It is however not the "silver bullet". When using CDISC standards end-to-end, other tools such as our SDTM-ETL$^{TM}$ software (there are also others) are much more user-friendly and effective. But also these "second generation" tools require that you know your study and that you understand SDTM.

## CDISC publishes new Controlled Terminology - also in ODM format

CDISC very recently updated all Controlled Terminology and published it on the NCI website.

This time also in CDISC ODM format (as ODM Codelists) - something that has been on our "wish list" for several years.

In earlier times, each time new Controlled Terminology was published, we needed to transform that into XML in order to be able to use it in our software packages. Though we had some programs for that, it was pretty painful as the Excel files in which the Controlled Terminology was published each time had a (sometimes even slightly) different formatting, meaning that we needed to adapt the programs each time.

But these times are now over!

The newly published Controlled Terminology replaces all earlier versions. The latest, complete versions thus are (with release date):

SDTM: 2011-06-10
SEND: 2011-06-07
ADaM: 2011-01-07
CDASH: 2011-04-08

Of course we will implement these codelists in all our software packages as soon as possible.

Special thanks are due to Lex Jansen (SAS) and Andy Fowler who made software available for automating the transformations from Excel to ODM.

The next step I would propose is that NCI also makes a web service available, e.g. allowing to send a request like "give me all controlled terms for VSPOS where VSTESTCD=DIABP".

## Cool Technology

This month it is not really about cool technology, but about "cool" clinical processes.

As the first of its kind, Pfizer has started a clinical trial in which there are no sites and no visits. The subjects participate from their homes using their PC and smartphones. Visits to a clinic are not planned except for a blood draw, which can however also be done at the subject's home. Medications will be mailed to the participants. They will keep diaries using smartphones, and will fill in a set of forms using a secure website four times throughout the study. The study is being overseen by a single group of doctors and nurses at the University of California, there are no other investigators or sites.

Although such a concept is of course not applicable to every clinical study, I think it is a major breakthrough. Very (positively) surprising is that the FDA approved all this, as their regulations are full of wordings like "site", "visit", "investigator". So "bravo" to the FDA to take this new path!

**Some snapshots taken from our ODM Study Designer implementing the new ODM-SDM extension: a study design with two "Drug treatment" arms and a "Chirurgy treatment" arm (top). Partial workflow for the LZZT study (bottom)**
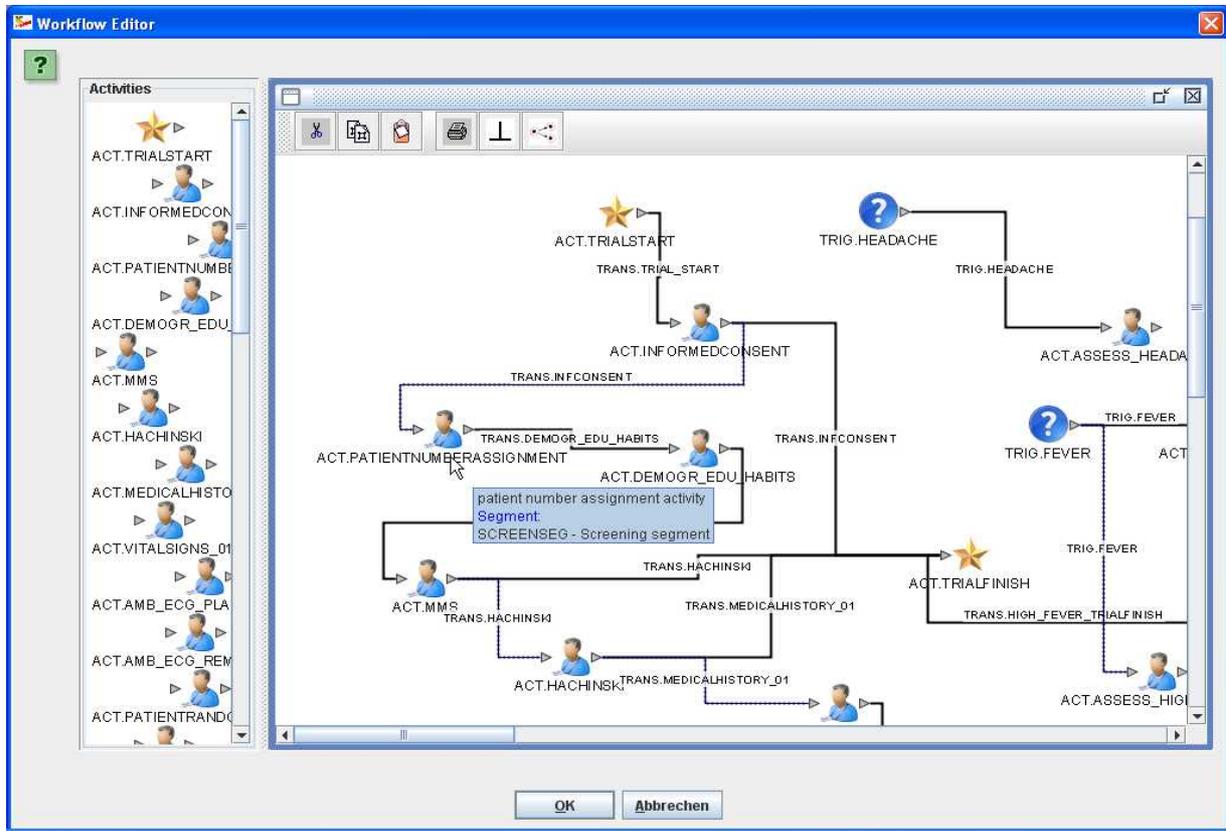
**The SMART Challenge: our app extracts the list of medications from the EHR/EMR (upper image) and uses them to prefill a CDASH form for Demographics or Concomitant Medications in the EDC system (center image).**
**After submission to the server an ODM file is generated (lower image)**

Clinical Research
Med List
Med Risk Maps
My App
MyDrugAdvice
NLP Meds
PinkHealth
Manage Apps
CHIP HMS © 2011

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 10 MG Oral Tablet | 30 | 2003-11-18 | 2007-08-03 | | 2007-09-02 | 1 | tablet | 1 | /d |
| rosuvastatin 20 MG Oral Tablet [Crestor] | 30 | 2005-09-10 | 2007-08-03 | | 2007-09-02 | 1 | tablet | 1 | /d |
| Hydrochlorothiazide 12.5 MG / Lisinopril 10 MG Oral Tablet | 30 | 2003-11-18 | 2007-09-02 | | 2007-10-02 | 1 | tablet | 1 | /d |
| rosuvastatin 20 MG Oral Tablet [Crestor] | 30 | 2005-09-10 | 2007-09-02 | | 2007-10-02 | 1 | tablet | 1 | /d |
| Hydrochlorothiazide 12.5 MG / Lisinopril 10 MG Oral Tablet | 30 | 2003-11-18 | 2007-10-06 | | 2007-11-05 | 1 | tablet | 1 | /d |
| rosuvastatin 10 MG Oral Tablet [Crestor] | 30 | 2005-01-24 | 2007-10-06 | | 2007-11-05 | 1 | tablet | 1 | /d |

**Patient is involved in Clinical Study # 123**

Transfer Medical Record data to CDASH Demographics CRF

Transfer Medical Record data to CDASH Concomitant Medications CRF

More information about the CDASH standard

troubleshooting

---

SMArt EMR

< Alex Lewis >                                     Jozef Aerts

SMART

Patient List
API Playground
Cardiac Risk
Med Adherence
Med Calendar
Med List
My App
Problems
Manage Apps

**Group: Concomittant medication**

[ Insert after selected Group ]   [ Remove selected Group ]

Medication No.*          1
Drug Name (Brand or Generic)*     Warfarin 5 MG Oral Tablet [Coumadin]

Total Dose Per Freqeuncy Checked*    1
Unit*            tablet
Frequency*        1/d
Start Date*       2008-07-30
Is medication still continuing*   No
Stop Date*        2008-10-28

Medication No.*          2
Drug Name (Brand or Generic)*     carvedilol 25 MG Oral Tablet

---

- <SubjectData SubjectKey="308194">
    <InvestigatorRef UserOID="Investigator"/>
  - <StudyEventData StudyEventOID="Concomitant Medications Visit">
    - <FormData FormOID="F_CM">
      - <ItemGroupData ItemGroupOID="IG_COMMON">
          <ItemData ItemOID="I_SITE" Value="CHB"/>
          <ItemData ItemOID="I_SUBJECTID" Value="308194"/>
          <ItemData ItemOID="I_VISIT" Value="2011-06-14"/>
          <ItemData ItemOID="I_VISITTIME" Value="16:28:10"/>
        </ItemGroupData>
      - <ItemGroupData ItemGroupOID="IG_CM_TAKEN">
          <ItemData ItemOID="I_CM_TAKEN" Value="1"/>
        </ItemGroupData>
      - <ItemGroupData ItemGroupOID="IG_CM" ItemGroupRepeatKey="1">
          <ItemData ItemOID="I_CM_NUMBER" Value="1"/>
          <ItemData ItemOID="I_CM_NAME" Value="Hydrochlorothiazide 12.5 MG / Lisinopril 10 MG Oral Tablet"/>
          <ItemData ItemOID="I_CM_DOSE" Value="1"/>
          <ItemData ItemOID="I_CM_DOSEUNIT" Value="tablet"/>
          <ItemData ItemOID="I_CM_DOSEFREQ" Value="1/d"/>

**Another app (from USGovXML.com) working on the same EHR: searching for relevant clinical studies on ClinicalTrials.gov**



---