PhUSE 2015

Paper CD04

Access to CDISC SHARE Metadata: A la carte, prix fixe, or table d'hôte?

Ken Stoltzfus, Accenture, Berwyn, USA

ABSTRACT

The CDISC SHARE (Shared Health And Clinical Research Electronic Library) repository is now available to help sponsors develop, integrate, and access the CDISC data standards. It is an innovative technology that is poised to change the way the industry utilizes standards metadata for clinical trials, advancing the industry's ability to efficiently create and analyze clinical data, and submit regulatory compliant data. The higher quality, standardized metadata provided by SHARE fosters benefits such as the promotion of process automation, data consistency, data re-use, and data aggregation.

Sponsor access to this standardized metadata is emerging as a vital ingredient to the full benefit of the CDISC SHARE tool. This paper intends to explore the challenges and potential solutions for sponsors regarding the current and future direction of the CDISC SHARE repository, and the industry intention for interactive access to it for the sponsor metadata management tools.

CDISC SHARE 101

SHARE is the global electronic metadata repository utilized by the CDISC organization to develop, manage, and provide access for the metadata that describes the CDISC standards in electronic format. It has been implemented to aid the users of CDISC standards and controlled terminologies in searching and leveraging the metadata about the standards, while also facilitating sponsors in their pursuit to integrate the CDISC standards into their own clinical processes in a manner which provides traceability of their clinical data throughout the clinical data lifecycle.

At a high-level, SHARE is currently composed of 2 main components, each with its own specific purpose:

- **iSHARE** - the interactive tool used by the CDISC standards development community for developing, governing, and publishing the standards. It is the source for content that is available for export in eSHARE.
- **eSHARE** - the eSHARE website serves subscribers that implement the CDISC standards, containing exports in multiple formats (e.g. ODM, Define-XML, CSV) as well as multiple versions of the standards. The eSHARE exports reflect the standards content being governed in iSHARE.

Interactive SHARE (**iSHARE**) is the cloud-based utility used by CDISC resources and its collaborative development community to interact with the standards metadata for purposes of developing, maintaining, and publishing it. iSHARE serves as an MDR for metadata management that is designed specifically for the CDISC standards. The standards development teams currently in progress producing new standards (CFAST, TAUGs, etc.) work closely with the SHARE Metadata Curators to construct new CDISC standards within the SHARE tool and subsequently publish the new standards to the user community at large. The iSHARE tool also provides tools for constructing and exporting the standards content in a number of machine-readable formats.

eSHARE is a website that maintains executable, downloadable links for each of the standards metadata published from iSHARE. Sponsors that are subscribers have easy, available access via the eSHARE website to the metadata in multiple formats and multiple versions, even having the option of receiving email notifications when new content is posted for download. As a general rule, preliminary work in-progress for upcoming new versions of a standard will not typically be published to eSHARE. This is to maintain the necessary critical step in the development process of "public comment" to ensure that the needs of the industry are being met. While the majority of eSHARE content currently are machine-readable versions of the CDISC standards metadata, there still do exist a few PDF versions for the near future. A recent example of a portion of the eSHARE catalog of accessible standards and formats is included below:

Date Posted	Content	Version	Type	Download Files
SDTMUG Virology				
2015-04-28	SDTM UG	1.0	For Review	Excel
SDTMIG Medical Devices 1.0				
2014-11-20	SDTM IG	1.0	Metadata	Define-XML v2.0 | Define-XML v1.0 | Excel
2014-11-20	SDTM IG	1.0	Document	PDF
SDTM 1.4 (IG 3.2)				
2015-02-02	SDTM IG Diff	3.2 - 3.1.3	For Review	Excel
2014-08-21	SDTM (IG)	1.4 (3.2)	Metadata	Define-XML v2.0 | Define-XML v1.0 | Excel
2014-08-21	SDTM + CT	1.4+2013-12-20	Bundle	Define-XML v2.0 | Excel
2014-08-21	SDTM (IG)	1.3 (3.1.3)	Document Bundle	PDF
SDTM 1.3 (IG 3.1.3)				
2014-08-21	SDTM (IG)	1.3 (3.1.3)	Metadata	Define-XML v2.0 | Define-XML v1.0 | Excel
2014-08-21	SDTM + CT	1.3+2013-12-20	Bundle	Define-XML v2.0 | Excel
2014-08-21	SDTM (IG)	1.3 (3.1.3)	Document Bundle	PDF
SDTM1.2 (IG 3.1.2)				
2014-06-02	SDTM (IG)	1.2 (3.1.2)	Metadata	Define-XML v2.0 | Define-XML v1.0 | CSV |Excel
2014-06-02	SDTM + CT	1.2+2013-12	Bundle	Define-XML v2.0 | Excel
2014-06-02	SDTM	1.2	Document	PDF
2014-06-	SDTM IG	3.1.2	Document	PDF

METADATA USE CASES

As regulatory authorities around the globe begin to demonstrate their intent to require the use of CDISC standards in regulatory submissions (see table below), sponsors are seeking a more consistent and efficient way to utilize the standards metadata for compliant submission data. The industry is seeing concrete dates on the future regulatory calendar which will require their data deliverables to be compliant.

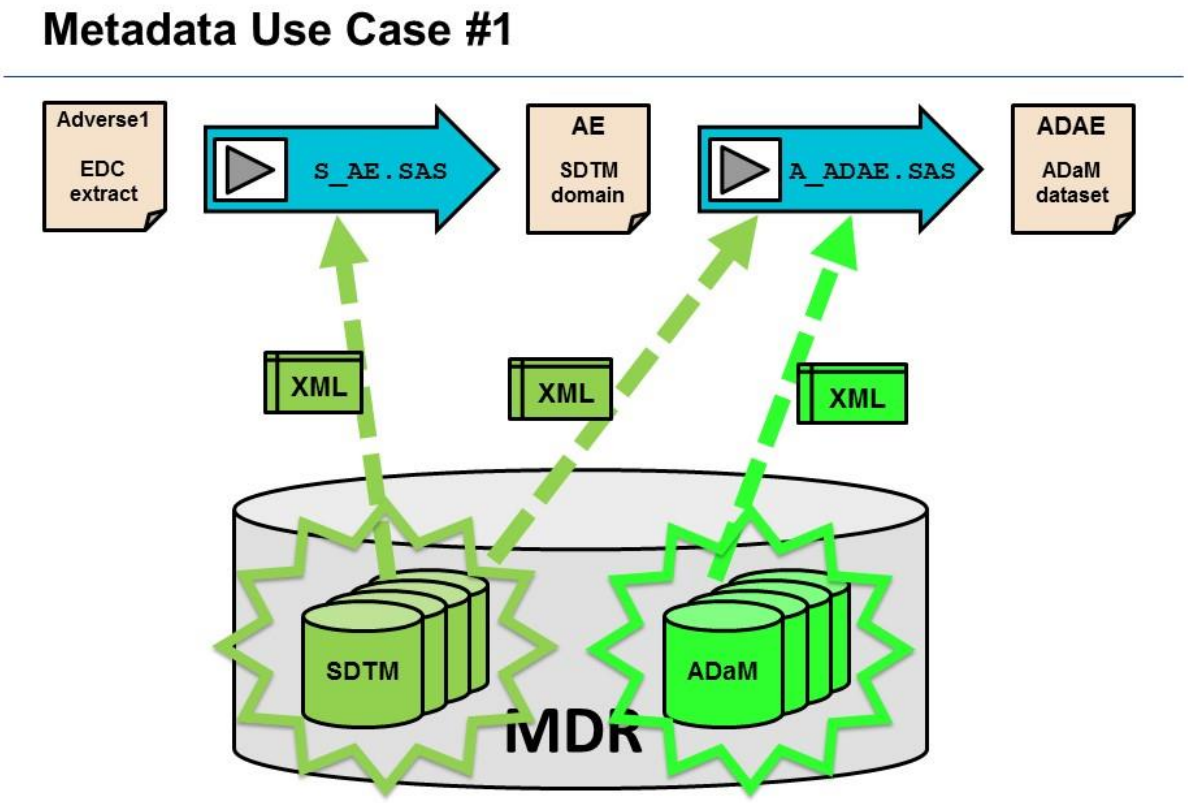
Region	Agency	Regulation	Key Date(s)
Japan	Pharmaceuticals Medical Devices Agency (PMDA)	***Basic Principles on Electronic Submission of Study Data for New Drug Applications*** Will require electronic clinical study data conforming to the CDISC SDTM and ADaM format [2]	Japan's FY-2016 (with a transitional period)
United States	Food and Drug Administration (FDA)	***Providing Regulatory Submissions in Electronic Format — Submissions Under Section 745A(a) of the Federal Food, Drug, and Cosmetic Act*** This legislation gives the FDA the ability to enforce CDISC standards[3] ***Providing Regulatory Submissions in Electronic Format — Standardized Study Data***[4] Clinical and non-clinical studies must use the CDISC standards in the Data Standards Catalog[5]	December 16, 2017 (Studies starting after this date must use the CDISC standards in the Data Standards Catalog)
European Union	European Medicines Agency (EMA)	Currently the European Medicines Agency Management Board on 12 June 2014 agreed the policy on publication of clinical trial data (***Publication and access to clinical-trial data, EMA/240810/2013***), together with more user-friendly amendments, and is currently in the process of being finalized.[1]	To Be Determined

The industry trend is beginning to include more and more sponsors going beyond just implementing the CDISC standards into their data deliverables, as they are actively seeking to deploy CDISC standards metadata as part of their clinical processes to automate the production of and make consistent, reproducible iterations of those regulatory-compliant data deliverables. With the clinical data lifecycle being the complex, multi-variate set of processes that it is, there is no one, single solution for where and how to best utilize this metadata. Each sponsor has its own set of variables within their own technical architecture that requires different solutions from sponsor to sponsor. Included in this section below are 3 potential processes that provide opportunities for leveraging standards metadata.

METADATA USE CASE #1

Numerous sponsors have already incorporated metadata from the SDTM and ADaM models into the production of the SDTM and ADaM datasets. Often this metadata is stored and maintained in MS Excel spreadsheets and has to be either manually reproduced into the dataset creation program or read as a spreadsheet directly by the program. In either scenario, there are challenges surrounding the integrity and consistency of the metadata coming out of these spreadsheets.

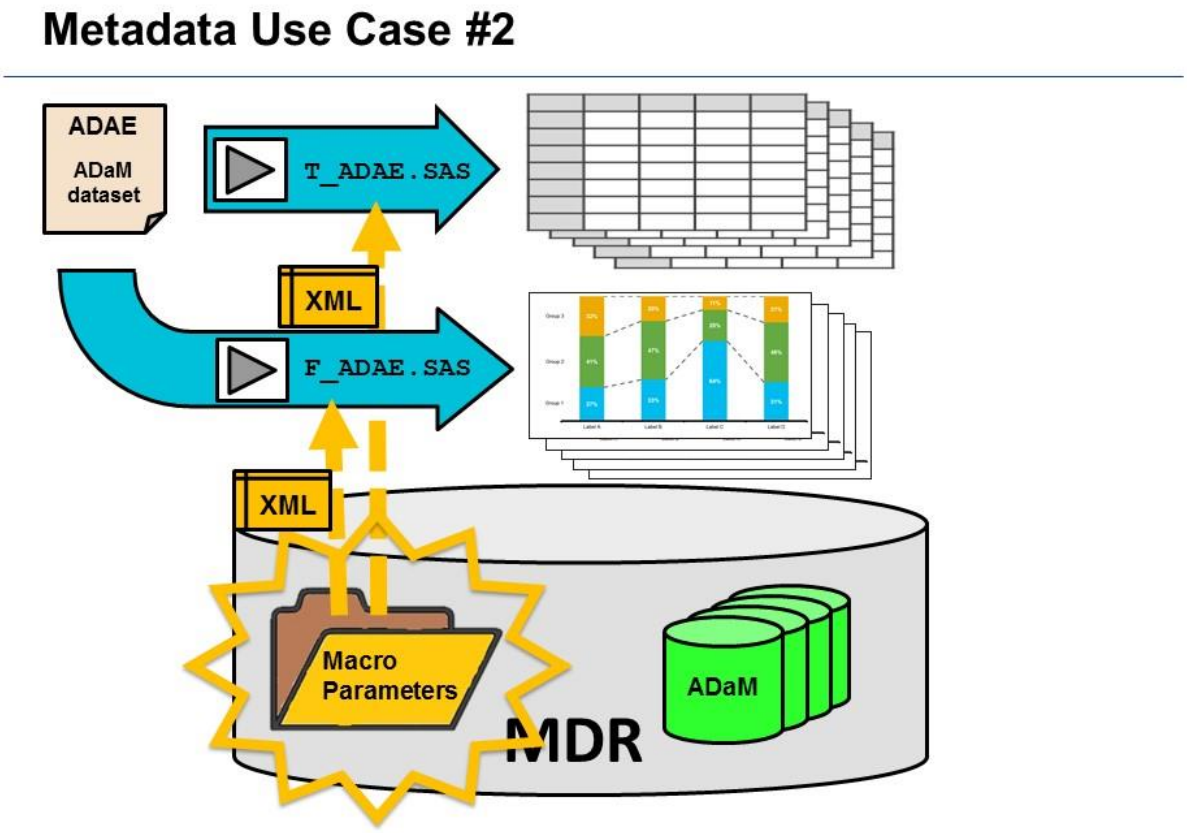
Potentially, SDTM and ADaM metadata stored in an MDR can provide the most recent, centrally managed, single version of the standards metadata directly into the dataset creation program (see diagram below).



METADATA USE CASE #2

An emerging set of standards metadata is results metadata, which can be utilized to prescribe the layout, structure, format, and breakdown of a sponsor's clinical results (tables, listing, and figures). Embedded in this results metadata can be reliable macro parameters that inform the table generator engine of the specific details of the construction of the output. These macro parameters can be enterprise-wide, TA-specific, or even study variations.

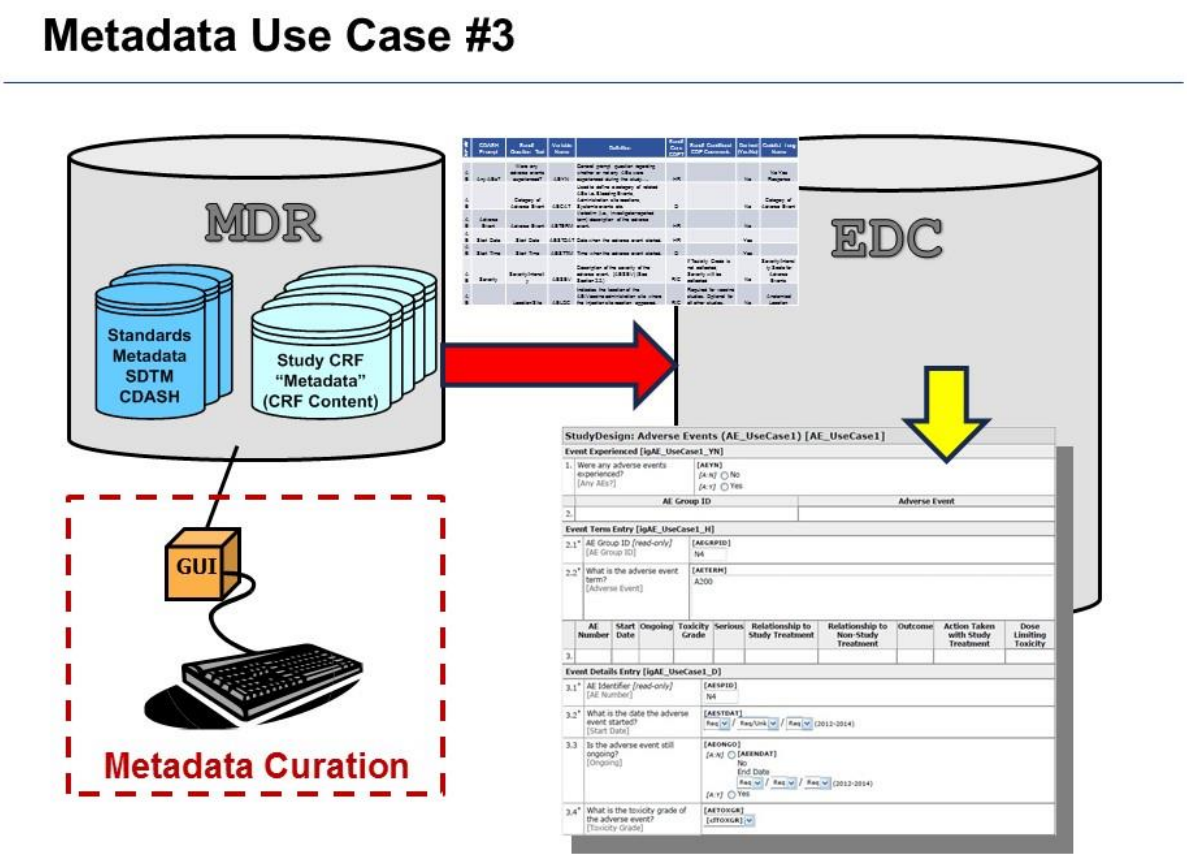
Potentially, these macro parameters can be stored in an MDR in order to be centrally managed and consistently applied across the sponsor's clinical programs. The export of these macro parameters from the MDR as standards metadata directly into the table generator engine ensures the deliberate use of the most recent, centrally managed, single version of the macro values (see diagram below).



METADATA USE CASE #3

As sponsors focus on the challenges to shorten the time from protocol approval to First Patient In (FPI), they are seeking alternate approaches to the often-laborious study build process. In some cases, they are utilizing the standards metadata that prescribes not only the collection data fields, but also the details of the eCRF in the EDC tool. This standards metadata is going beyond just describing the variable names, labels, lengths, etc., but also the attributes of the eCRF itself (question text, required usage, read-only, selection lists, etc.).

Potentially, this eCRF metadata can be stored in an MDR in order to be centrally managed to create a consistent look in front of the investigator while also maintaining a direct linkage to the regulatory-compliant variables. As the EDC system consumes the MDR export of this eCRF metadata, the opportunity for quickly producing library-selected, standards-based eCRFs can become a reality (see diagram below).



STARTING DOWN THE API PATH

While eSHARE is the near-term solution for the distribution of SHARE metadata, the future goal of both SHARE and the industry is for a more progressive approach that incorporates a more dynamic/complete interface between SHARE and the sponsor's metadata registry (MDR). CDISC and sponsors have quickly recognized that manual consumption of the SHARE metadata via a labor-intensive process is not a model that can be maintained efficiently over the long-term.

As the industry works more and more with standards metadata within their clinical processes, they are gaining the necessary experience to align with the CDISC SHARE team's vision of how the standards metadata can produce system interoperability[6] to fully maximize the use of their own MDR and the associated clinical processes.

- Increases process automation opportunities
- Increases metadata quality
- Improves data quality through standardization
- Improves data quality through semantics
- Encourages data re-use
- Facilitates data aggregation
- Improves data exchange across organizations
- Improves end-to-end data lifecycle efficiency

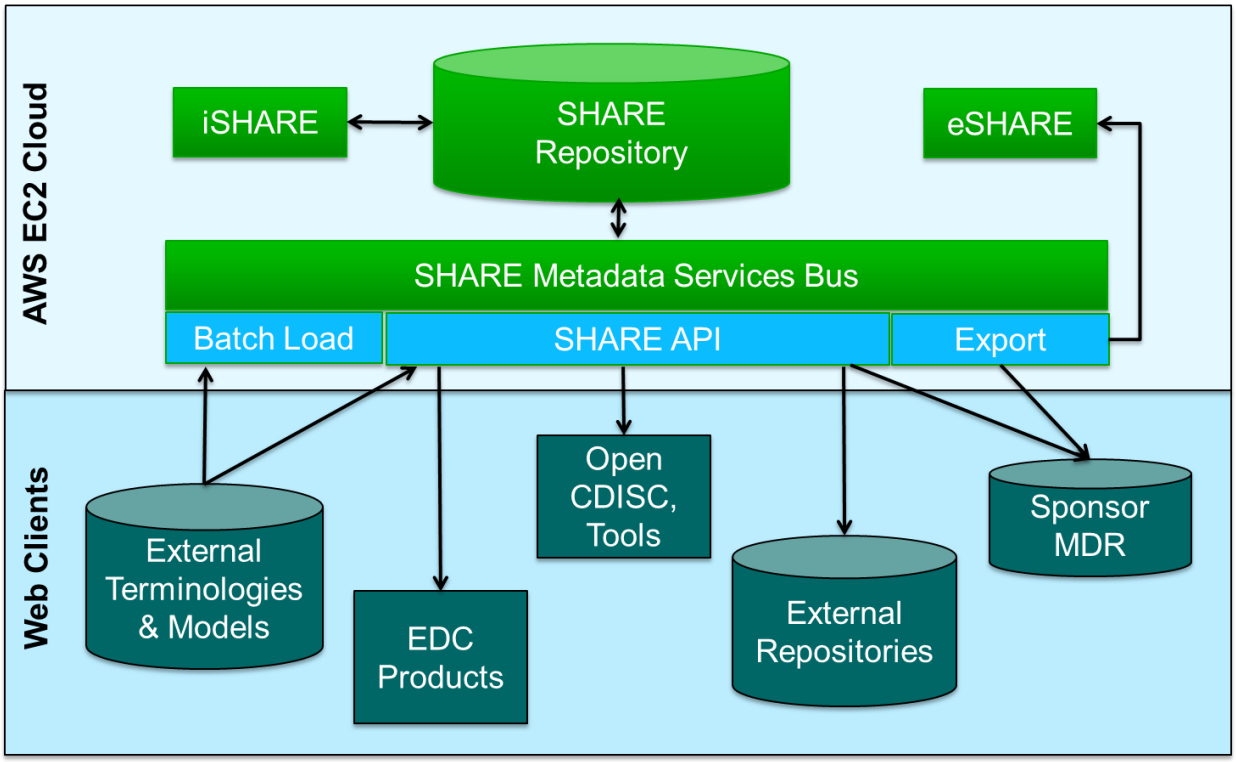
In 2015, the CDISC SHARE team started a pilot to enable automated programmatic access (API) to SHARE content from other systems, including other MDRs. The API will, for example, enable subscribers to identify content updates, allow interactive calls to retrieve targeted extracts of metadata, and enable automatic updates to standards metadata among sponsor MDRs. The pilot seeks to ensure that the API functionality will meet the most essential business needs to key stakeholders, including sponsor MDRs, commercial MDR providers, contract research organizations (CRO), other technology providers, and possibly other research organizations.

Pilot participant organizations currently include:

- **CDISC**
- Accenture
- Business and Decisions
- Cambridge Semantics
- Entimo
- eTriks
- Fujitsu

While the CDISC SHARE API Pilot is still in its early stages, its goals include enabling automated programmatic access to SHARE content, developing use cases, enhancing design aspects, and review and test the Web API features. It is also expected that the draft specification for the API that currently exists will be updated and delivered as a final specification during the pilot.

A preliminary high-level diagram of the CDISC SHARE API[6]:



CONCLUSION

The industry now stands on the precipice of how the CDISC SHARE repository will alter how sponsors will manage and incorporate CDISC standards into their clinical data processes. MDR tools will play an even more important role in helping sponsors utilize CDISC standards to provide successful regulatory-compliant submissions. The development of a CDISC SHARE interface that meets the business needs of the industry will be a significant catalyst in maximizing the role of the MDR.

REFERENCES

[1] Publication and access to clinical-trial data, EMA/240810/2013

[2] Basic Principles on Electronic Submission of Study Data for New Drug Applications, Evaluation and Licensing Division, Pharmaceutical and Food Safety Bureau, Japan Ministry of Health, Labour and Welfare (PFSB/ELD Notification No. 0620-6;published on June 20,2014)

[3] Providing Regulatory Submissions in Electronic Format — Submissions Under Section 745A(a) of the Federal Food, Drug, and Cosmetic Act, U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), and Center for Biologics Evaluation and Research (CBER) (February 2014)

[4] Providing Regulatory Submissions in Electronic Format — Standardized Study Data, U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), and Center for Biologics Evaluation and Research (CBER) (February 2014)

[5] FDA Data Standards Catalog (v3.0 20140117), available as a download from the Study Data Standards Resources webpage at http://www.fda.gov/forindustry/datastandards/studydatastandards/default.htm

[6] Using Information Standards to Drive eClinical Interoperability, Samuel Hume, DIA 2015.

ACKNOWLEDGMENTS

I wish to acknowledge my colleagues in the industry, Sam Hume, Barry Cohen, and Dave Evans, who have provided guidance and information regarding topics of metadata management and standards metadata.

CONTACT INFORMATION

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

Ken Stoltzfus
Accenture
Accelerated R&D Life Sciences - Clinical Data Strategies
1160 W. Swedesford Rd.
Berwyn, PA 19312 USA
Email: kenneth.m.stoltzfus@accenture.com