

### Multi-Sponsor Data Transparency: A Group Approach To Sharing

Bill Gibson, SAS Institute, Cary, NC, USA

#### ABSTRACT

In 2012, the idea of pharmaceutical companies providing their proprietary clinical trials data as a public resource would have met with skepticism and disbelief. Fast-forward to 2013 and 2014, and pharmaceutical companies are racing to market with implemented solutions that not only enable them to share their own company trial data, but enables that data to be combined with data from other pharmaceutical companies. Learn about one approach to data transparency designed and in use by multiple sponsor organizations. This approach enables researchers to view, analyze, and report on data from one or more studies from one or more sponsors. Main processes and options for sponsors (including data preparation, upload, review usage metrics) and researchers activities (access data from research grants, analyze data via SAS and R, request downloads of in-system developed reports) will be discussed in this paper.

#### INTRODUCTION

The life sciences industry seems to have come to agreement that greater access to patient-level data is a good thing – good for science, good for business and good for humanity. Data sharing can lead to discovery of new trends and associations that generate new insights or hypotheses for further research. Data sharing enables objective, third-party review and validation of study results, thereby building public trust. Data sharing honors the valuable information provided by patients and researchers in previous clinical trials and extends the future value of their efforts. Regulatory bodies like the European Medicines Authority are also developing regulations that will require a much greater sharing of clinical trial results and patient-level data.

But for all the benefits, there are caveats as well. Unless appropriate safeguards are in place, open access to clinical trial data could compromise patient privacy, enable faulty science, and be a resource-intensive burden for trial sponsors and data stewards. These are all outcomes the industry wants to avoid before going too far down this path. What information should be shared, with whom and for what purposes? How much demand is anticipated? How should this access and use be managed? How far back in time should study data be made available? How do you ensure patient privacy without unduly limiting the research value of the data? What should the information delivery and analysis platform look like? How should outputs from this open access be managed?

In February 2014, SAS brought together leaders from across the global pharmaceutical industry to share the latest updates on regulatory and technology fronts, stimulate conversation on the best way to move forward, and to see what others are doing – or not – to share their clinical trial data.

The spirit of the event marked an evolution in culture. Only a few months ago, the conversation might have produced nothing but arguments against data sharing. The mood has shifted to one of positive intention. It's just a matter of determining how to go about it.

#### THE MULTI-SPONSOR APPROACH

Many biopharmaceutical companies are electing to give researchers the option to access data from multiple sponsor companies within the same medical research project. This data transparency configuration is referred to as a multi-sponsor implementation.

The primary advantage of a multi-sponsor implementation for researchers is increased access to clinical trial data across multiple companies to allow greater scientific exploration, such as drug comparisons across sponsors in a single project. Other advantages include:

- A single, external interface across all sponsor companies.
- Clinical trial data that is segregated by company but accessible by researchers across all participating companies.
- A single process for researcher requests that can span multiple companies' data.
- A single, independent review board and process.
- Availability of all analytical and visualization tools within a secure environment.
- Researcher access to clinical trial data from many sponsors for use within a single research project. This allows for activities like meta-analysis of many therapies from multiple vendors that are used the same indication. This is something typically not possible from within a sponsor company.

# PhUSE 2014

- More rapid implementation of a standard solution with fewer implementation decisions to make, reducing time needed to reach “researcher-ready” as well as monetary costs.

## MAIN SPONSOR PROCESSES

There are many steps required to be completed by a data sharing sponsor before data can be provided to a researcher or research group. While this paper is focused on a specific data sharing environment and those processes, a number of foundational decisions also need to be decided by each sponsor organization. Those include what studies are to be potentially shared. Each sponsor needs to plan what studies can be made available, what counts as safe data to share (i.e. de-identification and the de-identification process), what group is responsible for authoring and maintaining the data sharing agreement, who will decide which research requests have merit and which do not, and what research request web site to use (the request portal and granting process are not part of the SAS data sharing solution).

The SAS Multi Sponsor Environment data sharing solution assumes SAS will provide the user, system, backup, and other IT administration functions. Each sponsor will need to notify SAS when a new research project has been approved along with each member of that research group and what studies are covered under that proposal (both the sponsor’s data as well as which other sponsors are also participating). Data will need to be prepared and uploaded into the sharing environment. Once data has been uploaded and made available to the appropriate researchers, a sponsor has a series of usage review and oversight tools to ensure their approved researchers are using the system for the intended purposes and to approve researcher requests for downloading of results and reports without participant identifiable information.

## DATA PREPARATION AND DE-IDENTIFICATION

Clinical trial data must be prepared prior to upload into the data sharing environment. This preparation process has likely already been defined within each sponsor organization. For the purposes of this paper, it is assumed that the data has been located, loaded, and participants have been de-identified according to the appropriate rules and regulations. There are currently no data standardization requirements (i.e. CDISC SDTM, MedDRA, etc.) for the data to be shared with researchers, but many sponsors using the MSE are intending to provide standardized data when feasible. Sponsors should consider the need to understand the provided data and probably aggregation challenges faced by researchers when preparing the data. The less data manipulation required by the researchers, the more transparent the data sharing will appear and thus the more time to focus on the science and analysis.

Once the data, its documentation and additional material for a study has been collected, de-identified, and prepared, it can be added to a ZIP or tar.gz file for uploading into the environment. Additionally, a study metadata file should be added to the root folder of the compressed archive. This metadata file is used to set the location, ownership, and read access to the data. This information can also be used to find the study for research requests made in the future as well to avoid repeated preparation and upload of the same trial data.

This is an example of the study metadata file:

```
# VALUES FOR TYPE ARE STUDY OR RESEARCH AREA
TYPE=STUDY
# VALUE FOR NAME MUST - MUST BE VALID DIRECTORY/FILE NAME
NAME=ACME-12345
# VALUE FOR DESCRIPTION TO BE USED BY RESEARCHER
DESCRIPTION=STUDY SAMPLE
SPONSORID=MSE_ACME
```

To update and existing uploaded study add:

```
APPENDFILES=YES
```

## INFORMATION UPLOAD

Uploading the study archive is performed through the administration web application. The upload portal make the process easy. All company-specific uploads made through the upload portal are logged for review. The portal page has an upload interface on the left and upload log on the right.

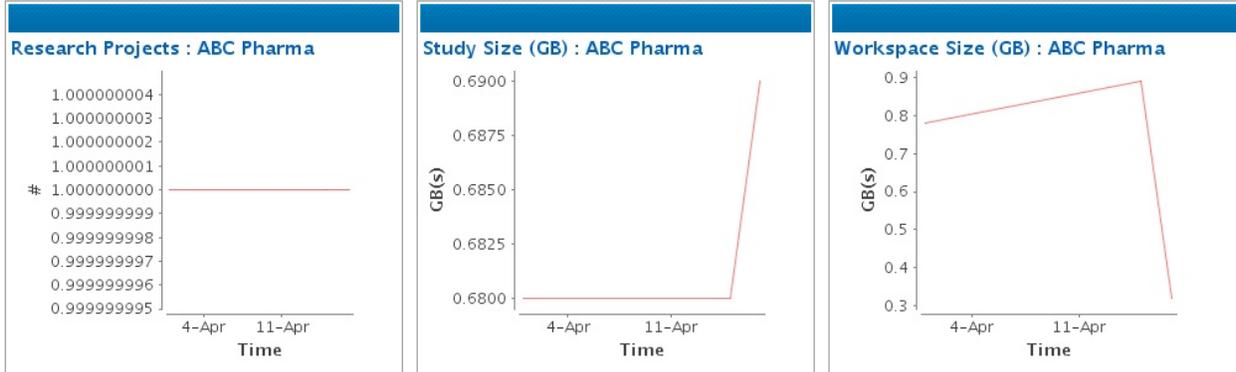
Feed File Summary:	
Current New Folder:	None
Current In Progress Folder:	None
Current Complete Folder:	<ul style="list-style-type: none"><li>• ACME-12345_WithSponsorID.zip_run_2014-04-09_14_49_46.log</li><li>• ACME-12987_WithSponsorID.zip_run_2014-04-09_14_24_04.log</li><li>• ACME-89765_WithSponsorID.zip_run_2014-04-09_14_44_37.log</li></ul>
Current Error Folder:	None

# PhUSE 2014

## USAGE METRICS REVIEW

Sponsor administrators can also review researcher activity, for the research projects approved by that sponsor. Researcher actions that are overseen by sponsors can be broken into two groups: data movement, and environment use. The data movement category includes export and export requests from the environment, as well as imports and import requests into the environment. For environment use, a summary of researcher access, content sizes, workspace storage, and tech support tracks are all reported via the sponsor administration web application.

Company: ABC Pharma Start Date: 02-MAR-2014 End Date: 01-MAY-2014 View



**User Access Summary: 4 record(s)**

Last Name	First Name	User ID	Role	Status (Request)	Status (Portal)	Status (Repo)	Company	Groups	Studies	Related Companies	History
Admin	ABC Pharma	cddt_labadmin	ABC Pharma Admin	Account Created	✓	✓	ABC Pharma	• CTD ABC Pharma Admin	• ABCP-85890 • ABCP-88877	• ABC Pharma	📘
Barnes	Karen	barnesk	Lead Researcher	Account Created	✓	✓	Pharma Research	• CTD RP-USERTEST - Full	• ABCP-85890 • ACME-12345 • ACME-12997 • CTD RP-USERTEST	• ABC Pharma • ACME	📘
Researcher	General	cddt_researcher	Researcher	Account Created	✓	✓	State U	• CTD RP-12345 - Full	• ABCP-85890 • ACME-12345 • ACME-12997 • CTD RP-12345	• ABC Pharma • ACME	📘
Researcher	Lead	cddt_leadresearcher	Lead Researcher	Account Created	✓	✓	State U	• CTD RP-12345 - Full	• ABCP-85890 • ACME-12345 • ACME-12997 • CTD RP-12345	• ABC Pharma • ACME	📘

**By User:**

Last Updated: 17-APR-14 12:58:46 EDT

User	User ID	Role	Organization	Size	Involved Companies
Researcher, Lead [LRES]	cddt_leadresearcher	Lead Researcher	State U	54.307365 MB	• ACME • ABC Pharma
Researcher, General [RES]	cddt_researcher	Researcher	State U	9.669193 MB	• ACME • ABC Pharma

**By Company:**

Last Updated: 17-APR-14 12:58:46 EDT

Company	Type	Status	Size
ABC Pharma	Large	Active	63.976562 MB

## Managing researcher download requests

One important aspect of the MSE is the ability to control data coming out of the environment (no downloads is a key requirement). Researchers will want to publish their results and findings, so some information does need to come out of the system. Sponsors or designates must review and approve download requests prior to a researcher accessing the results outside of the environment. While some requests can be filtered prior to sponsor review, some requests will need manual review. If the research project spans studies from multiple sponsors, each sponsor must approve the request. The request review and approval page provides status of the request as well as access to the requested archive to be downloaded.

# PhUSE 2014

Export Request:	
ID:	9
User:	Barnes, Karen [LRES]
Summary / Details:	test
Option:	Full - Latest File
Status:	Export Created / Approval Required for View
Deleted?	<input type="checkbox"/>
Revision:	6
Last Modified By:	System
Last Modified Date:	23-APR-14 11:03:52 EDT
Comments:	
<a href="#">Update</a>	

Export Paths:
• INCLUDE: /CTDT MSE/CTDT RP-USERTEST/Files/exports

Generated Exports:		
File	Date	Comments
<a href="#">exports_23-APR-14_16_01_47.zip</a>	23-APR-14 11:01:49 EDT System	Generated export, total files = 3

Company Approval(s):	
<a href="#">Update Approval</a>   <a href="#">Email Alert Companies</a>	
Company	Approval Status
ABC Pharma	Not Indicated
ACME	Not Indicated

User Information:	
Name:	Barnes, Karen [LRES]
User ID:	barnesk
Email:	kevin.alderton@sas.com
Role:	Lead Researcher
Company:	Pharma Research
Post or Position:	Lead Researcher
Phone:	011 44 1276 811888
Fax:	
Time Zone:	UK GMT/BST
Address:	

## History:

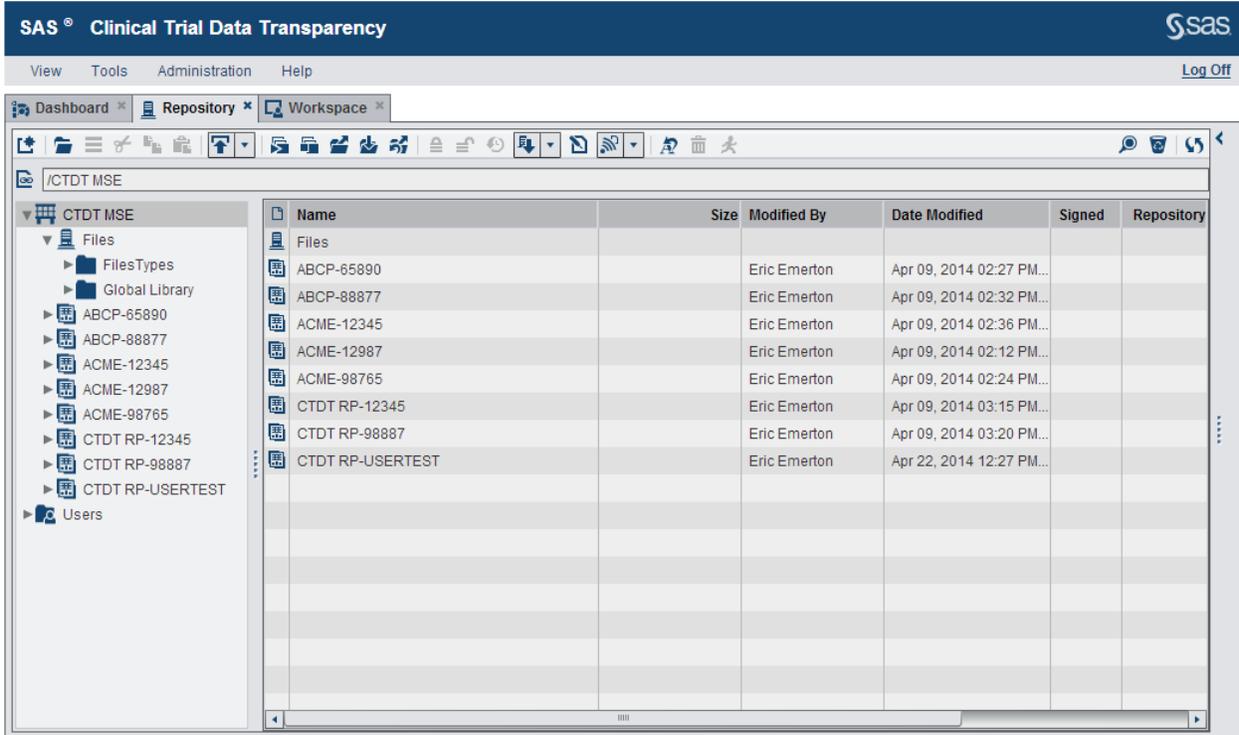
### RESEARCHER ACTIVITIES

The point of sharing data is to provide independent researchers a controlled way to access clinical trial data. The value to researchers is the ability to gain access to study data easily via a clear request path. Further, the additional value of the MSE is to provide data from many sponsors. While many of the initial research requests are to simply validate the sponsor published results, the greater purpose of sharing data is to enable the ability to find new insights not reasonably possible with single product, single sponsor sourced data. Goals beyond validation of published results includes examination of subset results (does this therapy work better or worse for a specific subset of patients), assessment of a specific outcome of interest (ex. Cardiovascular) – often adverse outcomes, deeper understanding of a disease and all available treatments as well as progress of trials and therapies over time (i.e. are we learning to design and run better trials, can we catch dropout candidates earlier or design trials with a lower risk of dropouts, or even is a placebo arm still needed).

### RESEARCH GRANTS, DATA ACCESS

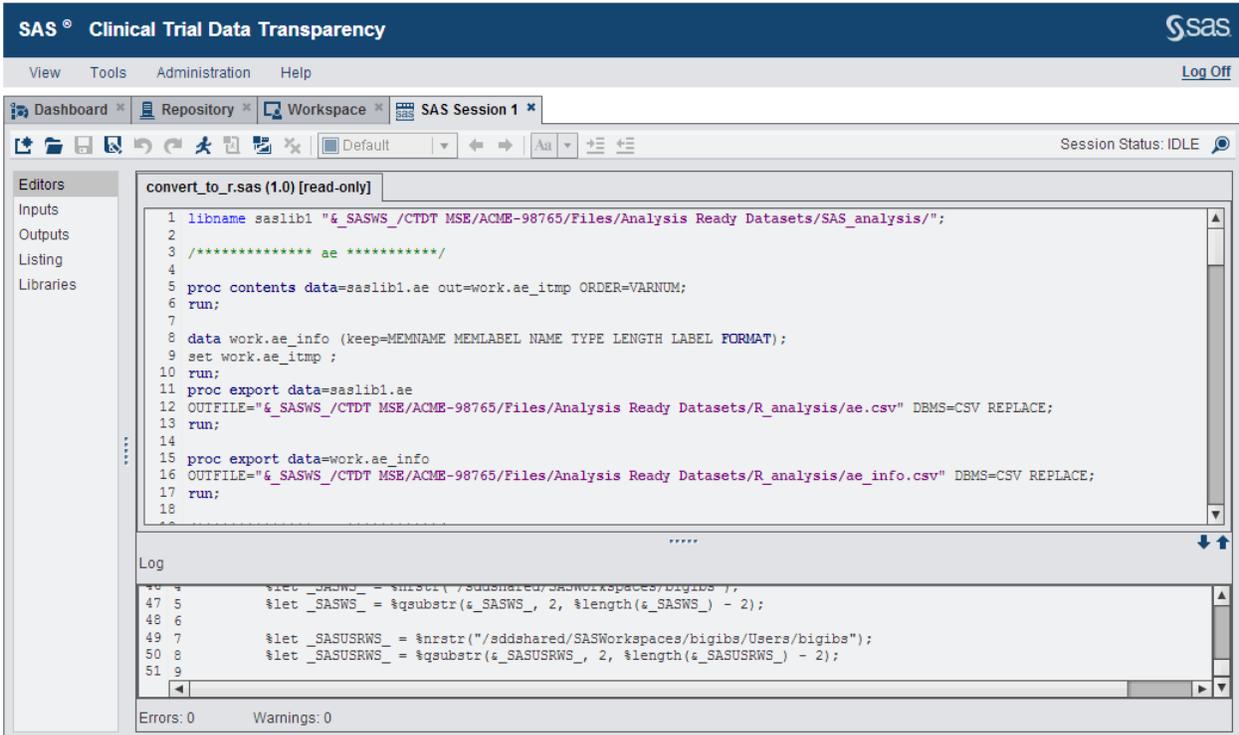
Once a research proposal has been approved by all sponsors listed in the research proposal and all sponsor required data use agreements have been signed, the SAS administration team creates the necessary changes to the MSE. The researchers are then able to connect to the MSE via web portal and given access to the researcher environment populated with appropriate tools for research and document authoring (includes the SAS Clinical Data Transparency repository, R Studio, OpenOffice, and Adobe Reader). This environment provides read-only access to the staged study data that was uploaded by the sponsor administrators as discussed in the prior section. This data can then be loaded into the SAS and R enabled development and analytics workbench. Multiple data and document viewers are provided so the researcher can review study documents explaining the available data and the study process as well as the ability to scan the data for additional familiarity. All available studies will appear in list with hierarchical content within the study root as shown in the example. The files area lists the content loaded and ready to use within the workspace area.

# PhUSE 2014



## DATA ANALYSIS VIA SAS AND R

The researcher can now prepare and analyze the data as per the research request. There are two analysis environments provided, SAS and R (via SAS IML and R Studio). The SAS programming environment is provided via a rich web-based client. It provides the typical color-coded editor, rolling log window, and output display. These programs, once ready to further work, can be moved into the version controlled area for the research group (if the research grant included more than once researcher) to collaborate and share results.



The R Studio programming interface is provided via the remote desktop environment. It has read and write access to the files in the researcher's workspace area. These files are also accessible from the SAS coding and repository

# PhUSE 2014

interface so completed or interim analysis programs can be shared with the rest of the research group. Data, programs, documents, and results are stored together and when ready can be packaged up for release from the environment.

## REPORT DOWNLOAD REQUEST

The lead researcher can request that package of files/results from completed or interim analyses be downloadable. This request process is simple and uses a similar web application as that used by sponsor administrators when uploading trial data.

The screenshot shows a web application interface for adding an import request. At the top, there is a navigation bar with 'Home' and 'CTDT Options'. Below this, a breadcrumb trail reads 'Section: CTDT Options | Exports | Imports' followed by an 'Options:' dropdown menu. On the right side of the breadcrumb trail, there are links: 'CTDT Options', 'Import List', and 'Add Import Request'. The main content area is titled 'Add Import Request' and contains the following elements: a message 'Files are added for import processing after add of the request.', a 'User:' field with the value 'Demo Researcher', an 'Option:' dropdown menu, and a 'Summary / Details:' text area. At the bottom right of the form, there is an 'Add' button.

Once a request is submitted it goes through the sponsor request approval process. This starts with running the contents of the package through a set of business rules for simple acceptance or rejection. For requests that do not meet the simple rules, each sponsor who contributed data to the research project must approve the request before the data is released to the researcher. In the future, this will likely evolve. A number of existing sponsors are looking to an independent 3<sup>rd</sup> party to approve these download requests as well as perform the initial research proposal requests as well.

## CONCLUSION

Only a few years ago, the possibility of many sponsor organizations sharing, through the same system, patient-level trial data in a formal way seemed impossible. Today this is a reality and one that is growing as more sponsor organizations are added to the MSE. Not only are researchers able to access de-identified patient data in a controlled environment, they are also able to use both SAS and R perform their planned analyses and research, all within one controlled environment. Eventually, even non-biopharma companies (medical researchers) are expected to also join as data providers as well as possible researchers.

## REFERENCES

SAS Institute, Inc. 2014. SAS Clinical Trial Data Transparency MSE Sponsor Admin Guide 1.0. Cary, NC: SAS Institute, Inc.

SAS Institute, Inc. 2014. SAS Clinical Trial Data Transparency MSE User Guide 1.0. Cary, NC: SAS Institute, Inc.

## RECOMMENDED READING

For more information about SAS Clinical Trial Data Transparency please visit [http://www.sas.com/en\\_us/industry/life-sciences.html#clinical-data-transparency](http://www.sas.com/en_us/industry/life-sciences.html#clinical-data-transparency).

## CONTACT INFORMATION

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

Bill Gibson  
SAS Institute  
SAS Campus Drive, R4465  
Cary, NC, 27513 USA  
Work Phone: +1-919-531-0996  
Email: [billj.gibson@sas.com](mailto:billj.gibson@sas.com)  
Web: [http://www.sas.com/en\\_us/industry/life-sciences.html](http://www.sas.com/en_us/industry/life-sciences.html)

Brand and product names are trademarks of their respective companies.