

## Using Workflows and Metadata Information to Standardize Business Processes in Pharmaceutical Programming

Peng Yang, Santen Inc., Emeryville, CA

Wei Liu, Santen Inc., Emeryville, CA

Julie Maddox, SAS, Cary, NC

### ABSTRACT

Pharmaceutical programming teams are getting increasingly more distributed and the tasks are becoming more complex. Standardizing business processes and ensuring consistency is crucial to the successful statistical analysis of clinical studies. Workflow templates can be defined to combine business process activities with the best practice techniques to be used as standards across an organization. Workflows can integrate both manual steps and complex processing steps that can use programmatic actions.

This paper describes how workflows can enable several common business processes in clinical data management, biostatistics and statistical programming. In addition, it demonstrates in technical detail how SAS<sup>®</sup> *Drug Development* implements these workflows. Possible application areas of workflows in the pharmaceutical programming will be discussed, such as workflows governing the development and validation of statistical programs using independent double programming practice, the interaction between statisticians and programming teams, and workflows for management of standards data exchange between a sponsor and the external partner (CRO). SAS<sup>®</sup> *Drug Development* also provides users Application Programming Interface (API) macros, which can be used to extract metadata information from SAS programs, logs and outputs. This enables the lead programmers or programming managers to visualize and manage many different aspects of the statistical programming projects. Finally, the workflow management system and metadata capability in SAS<sup>®</sup> *Drug Development* contributes to optimizing the described business processes and allowing management to make fact-based decisions, improve quality and efficiency of registration filings, and optimize the conduct of future clinical trials.

### INTRODUCTION

Workflow has been commonly used in document review and approval using document management software, whereas it has not been widely adopted in statistical programming processes. A workflow defines a sequence of tasks with assigned individuals to complete the tasks. Once an individual completes the assigned task, the item is automatically pushed to the next step until the entire collection of tasks has been completed. Implementing workflows in statistical programming area can help ensure programming activities are carried out in the correct sequence as defined by company Standard Operating Procedures (SOPs) or business guidelines. Workflows can also improve efficiency and productivity as the team members can focus on doing the work, rather than on managing the flow of tasks.

### WORKFLOWS IN PRACTICE

Below we describe three uses cases where workflows have been applied in statistical programming and clinical data management. All these use cases show how common business processes are translated into workflows and will be used later as examples to discuss more detailed aspects of workflows in SAS<sup>®</sup> *Drug Development*.

#### USE CASE 1: LIFE CYCLE OF A STATISTICAL PROGRAM

A rigorous life cycle of a statistical SAS program in a clinical study is deeply embedded in the practices of the clinical programming community. The example below illustrates the development and validation of statistical program which has inferential statistics. The programmer whose program outputs will be eventually used in production is the source programmer. Testing or validation can be performed with different methods, including independent double programming, code review, spot checks and cross checks. When the risk and complexity level is high, double programming is often used, in which a validation programmer independently produces the outputs with same specification. The validation programmer will check whether the source outputs and validation outputs are the same, whether the source output meet specification requirement and whether the log is free of error. If validation is failed, comments are sent back to the source programmer to make modification. After a program passes validation, it will be

moved on to a statistician for review. The statistician will check the inferential statistics portion and send comments if there are errors. Once the statistician is satisfied that the program produces the expected results, the program can be promoted and considered as part of the production environment or dry run whichever is applicable for the study (Figure 1). If no inferential statistics is needed in the program then the statistician quality control review step may not be included.

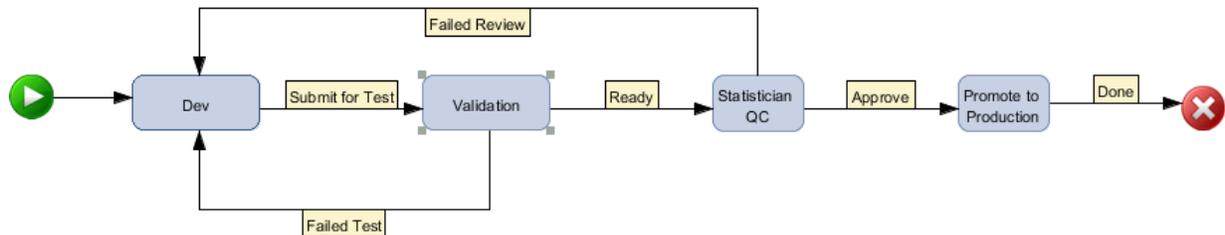


Figure 1: Workflow for statistical program validation cycle

### USE CASE 2: LIFE CYCLE OF PRODUCING FINAL STATISTICAL RESULTS FOR A DOUBLE BLINDED STUDY

Many clinical studies are double blinded in design; therefore it is useful to develop a process around un-blinding and final production run, which is critical to the success of generating study results. To produce top line results immediately after the database lock and un-blinding, a set of datasets and Tables, Figures and Listings (TFLs) are programmed with dummy treatment codes and generated in a dry run step. In the example below, we initiate the process with a dry run, where the study lead programmer runs all validated programs in batch. Team review is needed to get any final changes or comments to the Tables, Figures and Listings (TFLs) outputs before database lock. If programming is outsourced to a CRO, this step can be a review step conducted by the sponsor on the dry run outputs. Issues will be sent back to the programming team to resolve. When the database is locked and true treatment assignment data are applied, the lead programmer will manage the batch run for the entire set of datasets and TFLs. Often times, another round of team review is conducted to identify issues or request any ad hoc analysis. Finally, all programs are moved into production environment for final production run. Since this workflow captures very high level tasks, it does not display every validation steps, but every batch run should be accomplished by proper validation steps to ensure the run is successful and the logs are clean. The step of un-blinding also requires database ready to be locked, which may include several check points. A separate workflow can be developed to manage that.

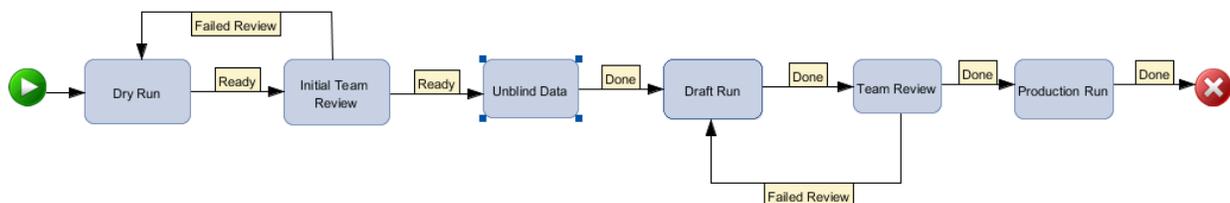


Figure 2: Workflow for Un-Blinding and Production Run

### USE CASE 3: LIFE CYCLE GOVERNING THE EXCHANGE OF CLINICAL DATA BETWEEN SPONSOR AND CRO

One of the transformations and changes clearly taking place in the industry is a renewed definition of clinical data operations, data management and submission-driven activities including collaboration with external organizations. Many pharmaceutical companies are responding to this new environment by setting up collaboration models with external clinical data management partners based on strong data standards governance processes. Figure 3 below, shows an example workflow where the incoming data (SDTM domains) and metadata (define.xml file or SAS tables containing the equivalent metadata) are used to compare against the organization's data standards, often based on CDISC (Clinical Data Interchange Standards Consortium) standard data models.<sup>1</sup> This workflow is an example of a more complex template because it consists of both user-based manual tasks and automated SAS job execution,

automated notifications, and it contains a decision point that leads to two different branches in the workflow. The initial step is for the CRO to upload the data into the study location. Once this task is completed, an automated SAS program is executed to validate whether the data meets the organization’s specifications. If the SAS program fails, then the CRO is automatically notified, and the CRO must make corrections and re-upload the data. If the SAS program passes, a notification is automatically sent to the project lead that the incoming data has passed the validation program.

This workflow allows pharmaceutical organizations to (partially) reduce their interactions with CRO’s in the context of specific trials and where pre-agreed data standards are used. It also produces a traceable and auditable record of the outcome of the comparison process.

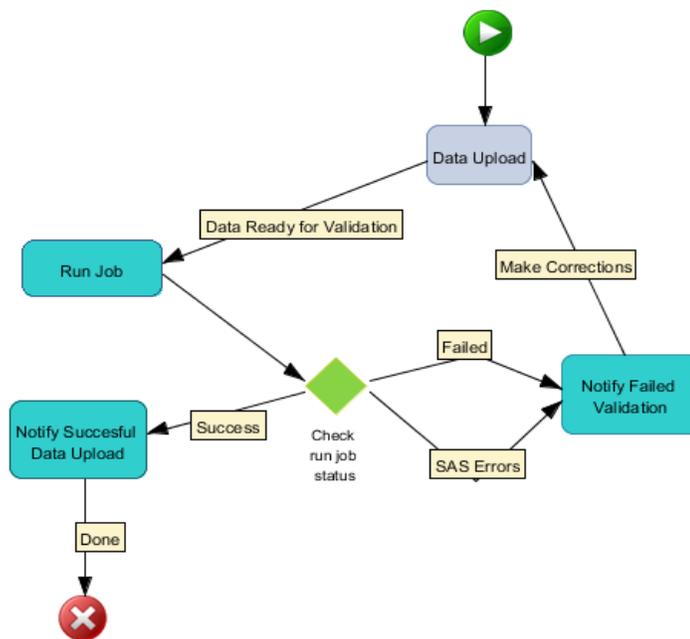


Figure 3: Workflow governing exchange of data between sponsor and CRO

## WORKFLOWS IN SAS® DRUG DEVELOPMENT

In SAS® Drug Development, the workflow management system surfaces workflows to the user as *work items*. A work item describes a unit of work that needs to be accomplished. For example, a work item can be defined to write code for a safety report, or tables, figures and listings program. A work item is defined using a workflow template which captures the tasks and flow describing a best practice. A work item can have due dates for tasks, task assignees, and can link into the specifications document of the task at hand.

Once a work item is activated in the system, the progress of the work item can be readily tracked. Project leads can re-assign tasks to other team members when the assigned resource is absent or if the workload becomes too heavy for the assigned resource. In addition, the flow of tasks through the work item is captured in the audit trail.

## WORK ITEM DEFINITION

For simple deliverables, the description field of the work item itself may be sufficient to detail the work to be accomplished. For more complex work items, attachments can be added providing more details, such as a program specification. Figure 4 below shows the General Properties for a new work item.



Figure 4: New Work Item General Properties

The next step in defining a work item is to select the appropriate workflow. The workflow selection determines the flow and order of the tasks for the work item. For example in Figure 1 above, the workflow describes the development, test and review process for moving a statistical program into a production clinical trial programming environment. This workflow has four tasks which need to be assigned. Figure 5 shows the Work Flow selection properties.

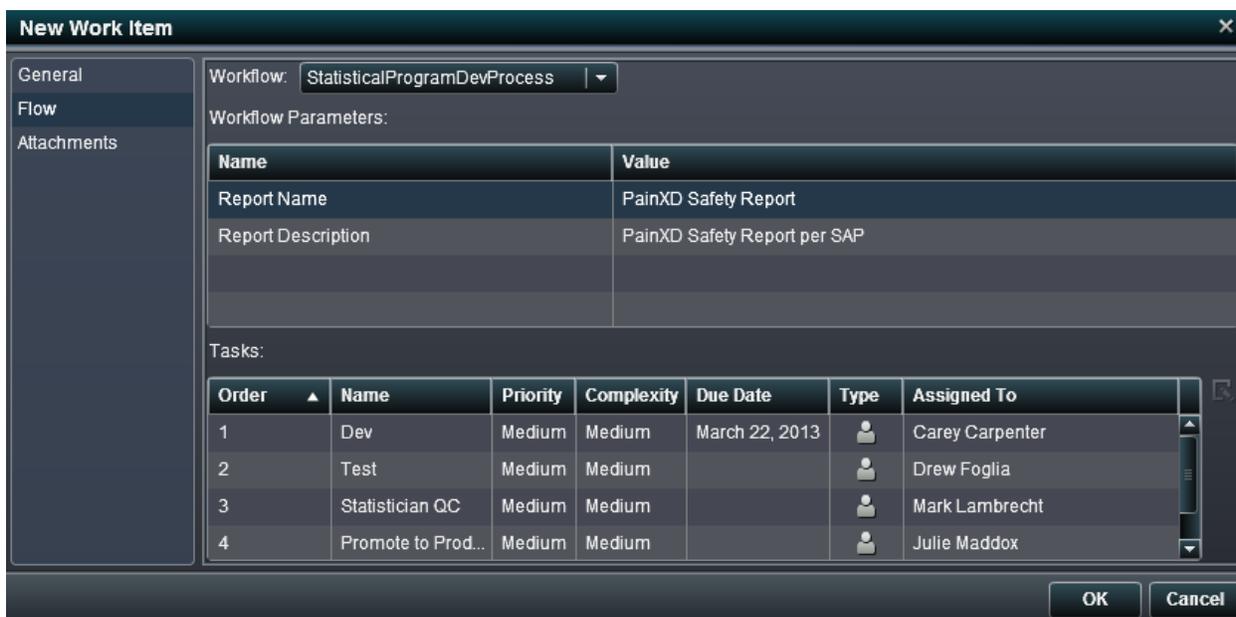


Figure 5: New Work Item Work Flow Properties

### WORK ITEM TASK ASSIGNMENT

Each of the tasks in a workflow can be defined as a User type task or as a Group type task. A task defined as a task type of User, is a task to be performed by an individual user. This type of task can be assigned to an individual or a group of users. When a User type task is assigned to a group, any individual within that group can claim the task to work on. For a Group type task, the task must be assigned to a group. In this case, any member of the group can work on the task without claiming it, and any member of the group can mark the task as complete. In the statistical program development workflow in Figure 1, the Promote to Production task might be defined as a Group task and assigned to a group of project leads. Any one of these leads can promote the program into production.

Figure 6 shows the properties of a task within a workflow. In the Task properties you can define the Due Date, the priority, complexity, and assign the task. Any task in a workflow which is not assigned will appear in the Task list of the project owner.

**Figure 6: Task Properties**

Once a work item is started, the tasks assigned to a particular user will appear in his individual Task List. Figure 7 below shows an individual’s task list, including the due dates. The user can select the work item link to open the link describing the work to be completed.

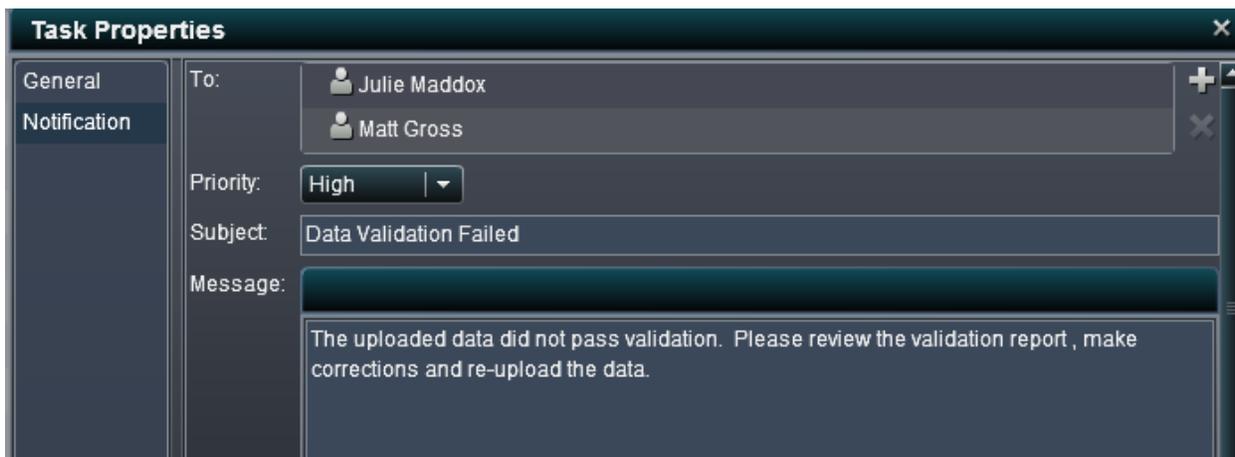
Status	Name	Work Item	Context	Due Date	Assigned To	Priority	Complexity
	Dev	<a href="#">Validate CRO Data</a>	PainXD	Apr 9, 2013	Julie Maddox	Medium	Medium
	Dev	<a href="#">ISE report</a>	PainXD	Mar 22, 2013	Julie Maddox	Medium	Medium
	Dev	<a href="#">ISS report</a>	PainXD	Mar 26, 2013	Julie Maddox	Medium	Medium

**Figure 7: Task List for an Individual**

## USING AUTOMATED TASKS IN WORKFLOWS

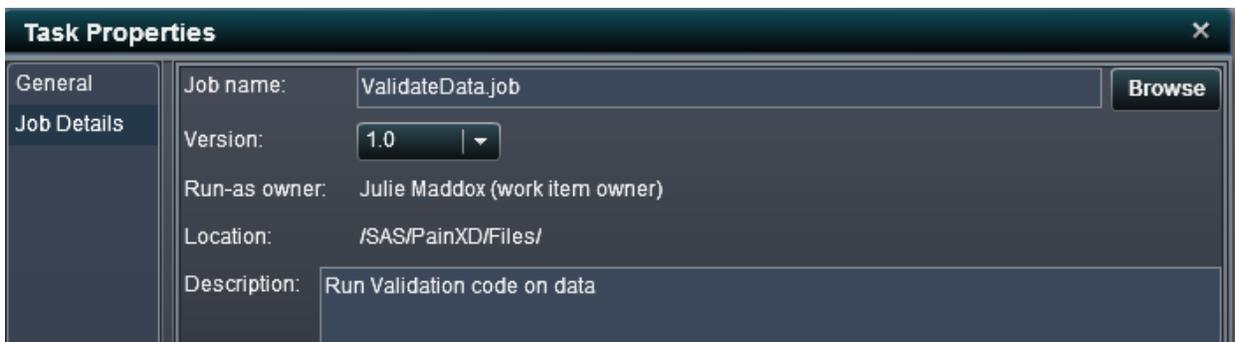
In addition to tasks which are performed manually by an individual, SAS® Drug Development provides two types of tasks which are automated. These automated tasks are a Notification task and an Execute Job task. Automated tasks are assigned to the system, and will not appear in the task list of any of the users included in the workflow. Figure 3 above illustrates how the automated notification and job execution tasks can be incorporated. In this example, the automated job task is executed and depending on the results of the execution an automated message is sent to the appropriate recipients.

Figure 8 below shows the properties for a notification task. The properties for this task include the list of users who should receive the notification as well as the content of the message.



**Figure 8: Automated Notification Task Properties**

Figure 9 below shows the properties of an execute job task. An Execute Job task will execute a SAS® Drug Development Job and return the execution status of the job to the workflow. The properties of this task include the name of the job to execute and any parameters to pass into the job.



**Figure 9: Automated Execute Job Task Properties**

## TRACKING PROGRESS

Once work items have been defined the project lead can track the progress of the work. Work items can have the following states: Started, Completed, Not Started, and Stopped. Figure 10 below shows the overall status of three work items associated with a specific project.

Status	Name	Context	Owner	Modified By	Date Modified
	Create DM report	PainXD	Julie Maddox	Julie Maddox	Mar 20, 2013 03:25 PM
	ISE report	PainXD	Julie Maddox	Julie Maddox	Mar 20, 2013 03:32 PM
	ISS report	PainXD	Julie Maddox	Julie Maddox	Mar 20, 2013 03:31 PM
	Validate CRO Data	PainXD	Julie Maddox	Julie Maddox	Mar 20, 2013 03:58 PM

**Figure 10: Work Item Status for Project**

To examine further details, the project lead can select the work item they are interested in, and drill into the list of tasks. Figure 11 shows the task flow for the selected work item. The active task and task assignees are displayed.

Tasks:							
Or.. ▲	Name	Active	Priority	Complexity	Due Date	Type	Assigned To
1	Dev		Medium	Medium	March 22, 2013		Julie Maddox
2	Test		Medium	Medium			Carey Carpenter
3	Statistician QC		Medium	Medium			Drew Foglia
4	Promote to Production		Medium	Medium			Julie Maddox

**Figure 11: Task Status for Work Item**

By selecting the Completed Tasks tab, the project lead can view the progression of the workflow. For example, in Figure 12 below, the comments show that the Dev task was repeated because a change to the report program was necessary after the Statistician QC review.

Name	Due Date	Assigned To	Priority	Complexity	Date Started	Date Ended	Completed Status	Comment
Dev	Mar 22, 2013	Julie Maddox	Medium	Medium	Mar 21, 2013..	Mar 21, 2013 ...	Submit for Test	Completed code
Test		Carey Carpenter	Medium	Medium	Mar 21, 2013..	Mar 21, 2013 ...	Ready	Validated code, no problems
Statistician QC		Drew Foglia	Medium	Medium	Mar 21, 2013..	Mar 21, 2013 ...	Failed Review	Need change formatting of report
Dev	Mar 22, 2013	Julie Maddox	Medium	Medium	Mar 21, 2013..	Mar 21, 2013 ...	Submit for Test	Updated formats

**Figure 12: Overview Status for Work Item**

## BENEFITS OF USING WORKFLOW

As described above, many programming tasks can be managed by using work flows. Table 1 summarizes the benefits from different perspective.

Programming Project Lead	Programming Team Members
Facilitates resource management - Resource can be centrally allocated into work items and project work can be divided in meaningful pieces	Get assignment from the system
Visualize project progression	Simplified communication as notification is streamlined
Audit trails for all work items	Higher level of compliance

**Table 1: Workflow benefits by role**

## STRATEGIES OF USING WORK FLOW AND WORK ITEMS IN SAS® DRUG DEVELOPMENT TO MANAGE STATISTICAL PROGRAMMING PROJECTS

Because the nature and the volume of a statistical programming project can vary a great deal, the strategies of implementing workflows and work items can also be quite different.

If the number of programs is not very big, work items can be created with one program per item. With this method, a program is moving on its own pace and it can be moved to the next step as soon as the current step is marked complete. On the other hand, if the volume is considerably large, work items can be created with a collection of programs; for instance, a project lead can create a work item of AE tables which can be a collection of 5 different AE table programs. The source programmer will only mark it complete when all of the programs in this work items are completed. This may lead to some increase in the wait time between source programmer and the validation programmer, but the benefits are the work is divided into a manageable number of logical groupings instead of excessive number of work items.

The statistical programming team structure also plays a big role in determining how the work flow is implemented. When programmers are co-located in the same office, the need of pushing every task through workflow may seem to be overhead, whereas a very large and distributed team with team members residing in different countries and multiple time zones will find the work flow process useful to ensure consistency and compliance.

## GATHERING INFORMATION FROM METADATA

SAS® Drug Development provides users with rich metadata, which can be utilized for programming project management purpose. Table 2 lists some metadata items and potential use in summarizing and analyzing project progress or process compliance.

Metadata	Possible ways to utilize the information
First creation date, last modified date of .sas	With these dates, duration of programming can be derived for individual programs or a collection of programs, such as the total time elapsed since the initiation of all SDTM programs.
Time stamp of .log and .lst	This can be used to document program run time. Project lead can use this to test whether programs are run in the correct order; for example, validation run should be after source run or tables run should be after the underlying datasets run.
User name for creation and modification	This can be used to summarize the work assignment by programmer.
isSigned	They can be used to summarize which files have been electronically signed.
isVersioned	Project lead can use this to check whether a set of programs have been put under version control.

**Table 2: SAS® Drug Development metadata and their potential usage**

## ACCESSING METADATA USING SAS® DRUG DEVELOPMENT API MACROS

How does user get access to metadata in SAS® Drug Development? One method is to call Application Programming Interface (API) macros. This requires logging in/out step using %SASDRUGDEV\_LOGIN(), %SASDRUGDEV\_LOGOUT(). Then %SASDRUGDEV\_CHILDREN() returns a SAS data set that contains the metadata for all of the objects within a container object in the SAS® Drug Development repository.

```
%SASDRUG_LOGIN(sdd_url=&SDD_URL,sdd_userid=&user_id, sdd_password=&sdd_pw);

%SASDRUGDEV_GETCHILDREN(sdd_path=&stdypath/sdtm,sdd_recursive=1,sas_dsname=sdtm);

%SASDRUGDEV_LOGOUT();
```

Once the dataset is obtained, the programming lead can use the information to create project level summaries, such as monitoring validation status using timestamp sequence business logic defined. For example, the project lead can look for situations where the source program is updated after the validation run. Traffic lighting techniques can be used to highlight the programs that are out of compliance. One example is shown in Figure 13 below.

Source Program Name	Source Program Completion Date	Source Program Run Date	Validation Program Name	Validation Program Run Date	Re-validation Required
dm.sas	2012-07-31T18:18:09	2012-08-01T16:13:14	v_dm.sas	2012-08-01T21:16:24	No
ae.sas	2012-08-02T21:30:27	2012-08-02T21:32:51	v_ae.sas	2012-08-01T21:16:24	Yes
adsl.sas	2012-08-02T21:30:27	2012-08-02T21:32:51	v_adsl.sas		Yes

**Figure 13: Metadata returned by API macro calls are used to generate summary spreadsheets**

## CONCLUSION

This paper discussed the benefits of using workflow and metadata management to increase the efficiency and compliance level in pharmaceutical programming projects. Using workflows, *SAS<sup>®</sup> Drug Development* provides the ability to implement, manage and track the progress of the programming projects. In addition, management reports can be readily produced from the metadata using the *SAS<sup>®</sup> Drug Development* API macros. Project team leads can use these tracking tools to correct errors on non-compliance and ensure high quality delivery of statistical analysis for registration filings.

## REFERENCES

1. SAS Drug Development as a platform governing data standards specification and clinical data exchange between pharmaceutical organizations, Mark Lambrecht and Peter Van Reusel, PhUSE 2011, Brighton, UK.

## ACKNOWLEDGMENTS

We would like to acknowledge Janardhan Ammisetti for his contributions in the development of project management tools using SDD API macros.

## CONTACT INFORMATION

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

Peng Yang  
Santen Inc.  
2100 Powell St  
Emeryville, CA 94608  
Work Phone: 415-900-6064  
E-mail: [pyang@santeninc.com](mailto:pyang@santeninc.com)  
Web: <http://www.santeninc.com>

Julie Maddox  
SAS Institute  
SAS Campus Drive  
Cary, CA 27519  
Work Phone: 919-531-6922  
E-mail: [julie.maddox@sas.com](mailto:julie.maddox@sas.com)  
Web: <http://www.sas.com>

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.