Paper #117-30

Designing a Global SAS®9 Architecture for Clinical Trials Analysis and Reporting

Greg Nelson - ThotWave Technologies, LLC. – Cary, NC
Michael Walega – Covance, Inc. – Princeton, NJ

ABSTRACT

Covance has a long tradition of utilizing SAS to analyze pre-clinical and clinical trails data, as drug candidates move from early discovery to late stage development. To continue to add value to shareholders, Covance has constantly looked to identify ways to fully leverage its global workforce and to maximize hardware and software toolsets. The goal: deliver high quality clinical trials reporting and analytic results to clients in an efficient and timely manner. Covance commissioned a study to evaluate their business processes relating to how they utilize SAS, to identify better ways to bring value to their clients, while keeping an eye on how SAS®9 could solve problems of today and for the future.

This presentation outlines the challenges that Covance has faced, along with a high level overview of what was done to create a new global SAS architecture that will result in a balanced, upgraded and technologically competitive SAS environment for Covance. Specific attention will be paid to the opportunities and trade-offs that presented themselves as part of the evaluation, and how the project will facilitate global resource sharing and implementation of global best practices. The result will be improved efficiency and productivity, and will provide greatly enhanced service options to Covance clients.

INTRODUCTION

Covance is a large clinical research organization (CRO) providing a variety of services to Biotechnology and Pharmaceutical organizations across the world. As such, Covance has a large investment in SAS software – not only from the perspective of the license fees, but a commitment to SAS in delivering state of the art techniques for reporting and analytics. With over one hundred SAS programmers and statisticians relying on SAS software to conduct activities for client projects, Covance recognized the opportunity to migrate from SAS Version 8.2 to SAS 9 as a chance not only to update its technology platform, but also to reevaluate and augment business processes to reflect the enhancements in a way that the technology could be used around the world.

Some of the capabilities and processes we hope to implement in the upgraded SAS environment include:

- The ability to fully leverage our global teams to work on extremely targeted problems – efficiently moving analytical activities to available resources rather than moving the resources to the location of the work
- Exploiting the numerous tools within SAS 9 to reduce keystroke programming effort
- More effectively utilize tools that are reusable at all levels of analysis and reporting

- Leverage electronically stored information within the clinical trials reporting and analysis world that will reduce the need for multiple touches (and thus multiple points for failure)

- Maintain a program development environment that enforces change control

Achieving these goals will allow us to meet a requirement we face today, and will face with increasing frequency in the future: how to effectively distribute our tasks to skilled Biometrics staff throughout the Covance world. It is clear that we are not simply talking about installing a new version of SAS, but rather a creating and re-tooling number of business processes that are highly intertwined with the technology that the Biometrics staff use to perform their jobs. In the next section, we will explore how SAS is used at Covance and how this scenario can be deconstructed to create a roadmap for Covance.

BUSINESS NEED DEFINED

Current State (How SAS is used today)

Within the Late Stage Development Services (LSDS) group at Covance, SAS is primarily used for analytical programming for Phase II-IV clinical studies, for other types of statistical studies, and for data management support. Figure 1 illustrates the typical data flow for a statistical study, clinical or otherwise. For clinical studies, the main data source is the Covance global instance of Oracle Clinical, but the majority of studies pull data from multiple sources. Studies often organize their raw data libraries to segregate data by source. Study outputs are typically statistical summary tables, graphical displays, and patient data listings (TFLs). Within LSDS the standard is to create derived analysis data sets as an intermediate step between raw data and study outputs. Using this approach, tabular output is most often generated by PROC REPORT and contains minimal computations. The final study outputs may be paper, electronic documents (e.g., Word, PDF), or data sets. Some variations that are highlighted on the data flow diagram are:

- Some studies pool data from multiple previous studies, typically in support of product registration applications.

- Derived data sets are not used in all cases.

- Some clients specify derived data set formats, either because they wish to receive these as study outputs or because they specify that TFLs must be produced from these derived data sets.

- Some clients receive raw data as a study output.

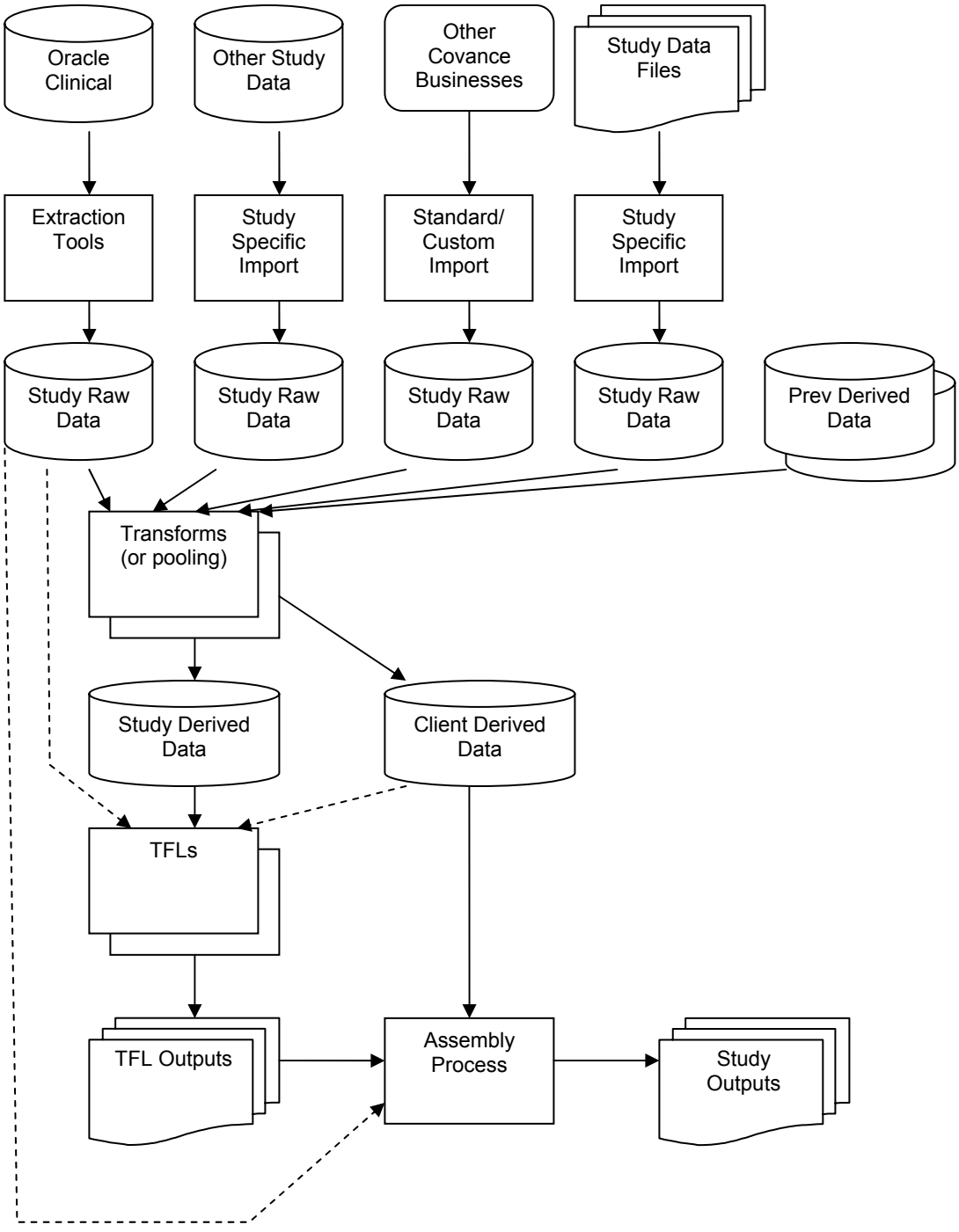

Figure 1. Study Data Flow

Within other Business Units of Covance, SAS is also used to conduct statistical analysis activities. The general flow of tasks is similar, but the major points of variation are:

- Some studies are very short in duration (e.g., one day turnaround).

- Some analyses do not employ clinical trials data.
- In comparison to LSDS, some groups are required to be more open-ended in the types of statistical analyses that are conducted, while others are more structured.

In addition, there is use of SAS for data management activities across Covance. These activities include creating SAS data sets and transport files, and general reporting/data quality check production.

Challenges

As a global organization, there are a number of challenges faced by Covance with regards to SAS software that required attention. These included global licensing, communication, and collaboration, and client demands for features found in SAS 9. These are summarized here:

- SAS is currently licensed on a site by site basis, making it difficult to assess the total cost of the software.
- Hardware limitations.
- Inconsistencies in technology (and certain processes as a result of this lack of consistency) resulting in inefficient use of resources.
- IBM-punch card key stroking rather than Object-oriented, drag-and-drop programming – the way people develop (non-interactive) and run/ assemble reports.
- Implementation of hot fixes (test, validate) and new versions of SAS is difficult to achieve globally.
- Not achieving the level of tool reuse that people feel should be possible.

Figure 2 summarizes the activities that are involved with LSDS statistical studies, along with challenges in the current SAS architecture and processes.

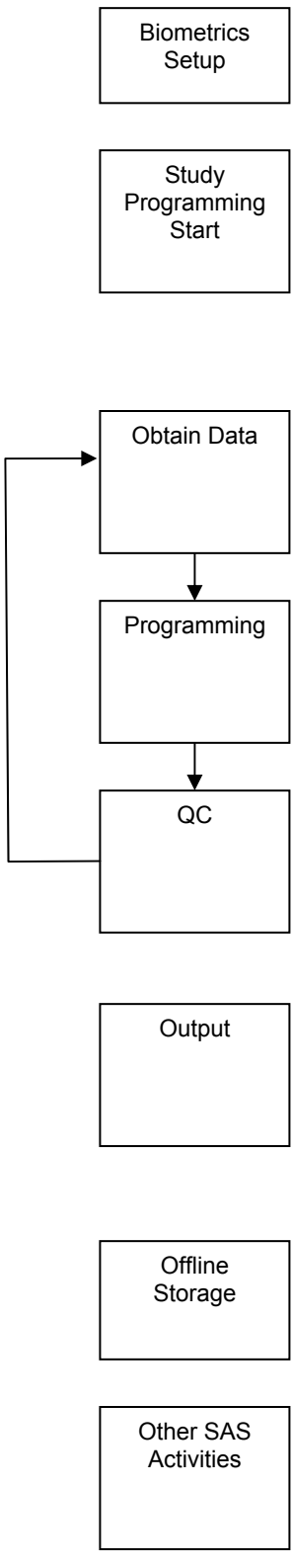

Figure 2. Study Activities

Business Process Improvement Opportunities

Taken as a whole, the future state of a global SAS architecture for LSDS should address the following goals:

- **Address current operational challenges.** Hardware capacity and throughput issues with the current SAS architecture,
- **Enable global resource management.** While LSDS is global in reach, the efficiency of sharing the workload between sites can be improved. Instead of requiring people to go to the work, the new architecture should allow work to go to people, wherever they are available.
- **Support the IT consolidation initiative.** Closely related is the goal of an architecture that is consistent with the data center consolidation strategy.
- **Centralize the licensing for SAS.** Begin by getting LSDS onto a common SAS architecture and licensing arrangement.
- **Institute business process improvements.** Putting new SAS tools in place provides the opportunity to implement improved processes at the same time, such as streamlining the import of study data and production of study outputs.
- **Facilitate shared knowledge and best practices.** The new architecture should be based on current best practices and encourage sharing new knowledge and learning.
- **Maintain and improve systems in place.** Several of the existing systems provide value to their constituency as they exist today (i.e., the business is currently operating). Their application functionality must be sustained and continually improved as it represents a functional state of the user's requirements today.
- **Define a path for the future.** The global SAS architecture project represents a solid vision for integrating sites into a single, global community. The sense of urgency to conform information assets is very high at the senior management level, as cross-organization sharing and utilization of resources are not possible until this platform and the associated business processes are integrated.

RECOMMENDATIONS

The requirements for the global SAS architecture are detailed in a Requirements Traceability Matrix which not only identifies the requirements, but categorizes and prioritizes them. For purposes of discussion, this section contains a summary of the desired characteristics of the architecture:

- Users (programmers/ statisticians) can work in a SAS environment where they are highly productive. Users should be able to interactively submit all or portions of a given program file, review the log, and review WORK and permanent data sets.

- Data should be made available to the user seamlessly. If the architecture involves multiple servers, data and compute services should reside in the location where it makes the most sense, without a lot of explicit work by users to move files.

- Code should be managed in an environment that is easy to update and change, but versioned centrally so that a history of changes is available. Access to projects should be explicitly granted.

- Interim output should be accessible in an environment that is easy to use (display, print, email).

- Importing of data should be easy, but programs and quality control steps retained for proof and imported data accessible to anyone working on the study.

- Metadata about studies should be leveraged throughout the analysis and reporting lifecycle, so that common information is not repetitively entered and maintained.

- Producing final output should be easier than it is today and better leverage the capability of SAS to create publishing output formats such as RTF and PDF.

- Final output (delivered to the client) should be versioned centrally.

- The project directory structure – including where SAS programs, quality control, data sets, macros, and pointers to global macros and data are located – should be more consistent across the organization to enable global sharing and reuse.

- The system should support communication and collaboration among a global team.

Possible Courses of Action

During this study, we considered a number of potential strategies for how to serve the Covance LSDS user community while considering the larger technical and organizational environments. At a high level, the possible courses of action were:

1. Do nothing (status quo option).

2. Migrate to a new platform (SAS9 and new hardware) while maintaining the same divergent business processes at each site (business continuity option).

3. Create a new architecture and supporting global business processes.

Do nothing

The option of doing nothing is highly risky. Current LSDS UNIX servers are not viable for a new, consolidated environment. LSDS users work in batch submit and interactive modes, and suffer from data interchange challenges resulting from the use of tools that do not import/export data well from the Windows platform.

Certainly this approach does not attempt to implement a global SAS architecture for LSDS or Covance. These challenges and unrealized benefits make this scenario highly undesirable.

Migrate Technology

A second option that was considered was to upgrade the hardware to a modern environment and SAS to Version 9, while using SAS in the same way that it is being used today. The advantage of this option is that huge improvements could be expected just by upgrading the hardware, particularly for a subset of the users that were only exposed to UNIX via command line. The expanded limits of disk space and processor capacity would show an immediate return on investment. Programs would run faster and there would be ample disk space to support the user community. By upgrading to SAS9, the user community would be able to take advantage of new functionality in Base SAS (new function calls), ODS with improved support for page numbering and output destinations, and new graphics capabilities including the new statistical graphics in V9.

While there is substantial value in stabilizing the SAS platform, the business continuity option does not attempt to create a consistent global approach for SAS usage. Thus, it does not deliver the benefits in global workload sharing that require a global system and process changes. Similarly, because processes would still be divergent, there would be continued inefficiencies in licensing SAS, maintaining the environments, tool development, deployment, and maintenance, and being unable to efficiently share reusable tools. Those sites that employ Windows SAS would still be limited to the processing capability of their workstations. UNIX SAS users would not realize some SAS9 benefits without making process changes. Finally, from a strategic IT vision perspective this is a significant financial investment without furthering the goal of data center consolidation.

Recommended Approach

As we considered the current state of SAS usage within Covance and the vision for a global SAS architecture, it became clear to us that a migration from the existing hardware and software to a robust, enterprise level, production SAS platform was warranted. The key characteristic of this new architecture is that it is global. This is discussed more fully in the remaining sections.

SAS 9 versus Previous Versions of SAS

The SAS9 architecture is fundamentally different from any prior version of SAS. In this architecture, SAS relies on a new component, the Metadata Server, to provide an information layer between the programs and the data they access. Metadata such as security permissions for SAS libraries and where the various SAS servers are running are maintained in a common repository.

By providing a single point of access for this kind of information, SAS servers can be located just about anywhere (consolidated or distributed) and on any platform (operating system and hardware), to be accessed by SAS clients. The SAS client tools now take on the form of special purpose applications designed to satisfy the needs of various types of SAS users. Management of SAS metadata is done through the SAS Management Console and Information Map Studio. Management of data is done through ETL Studio in addition to the existing SAS

products such as SAS/Access and Base SAS. Tools used to access information now include a substantially improved Enterprise Guide, the new add-in for Microsoft Office, and the web-based clients (Information Delivery Portal and Web Report Studio), as well as Base SAS components.

The SAS9 platform can therefore be configured in any number of ways, including a myriad of server topologies and client offerings. We considered five possible profiles of SAS products to realize a development environment for the generation of statistical summary tables, graphical displays, and patient data listings: Multi-tiered, PC SAS with file sharing, distributed environment using "classic" SAS connectivity, thin client, and hybrid.

These profiles make use of the new services in SAS9 in varying degrees. There are benefits and drawbacks to each architecture that were carefully considered for their business impact – both positive and negative – on the value added to Covance, implementation, efficient usage, and ease of maintenance.

The options that were considered are not necessarily mutually exclusive. Figure 3 depicts some ideas about how they might be combined.

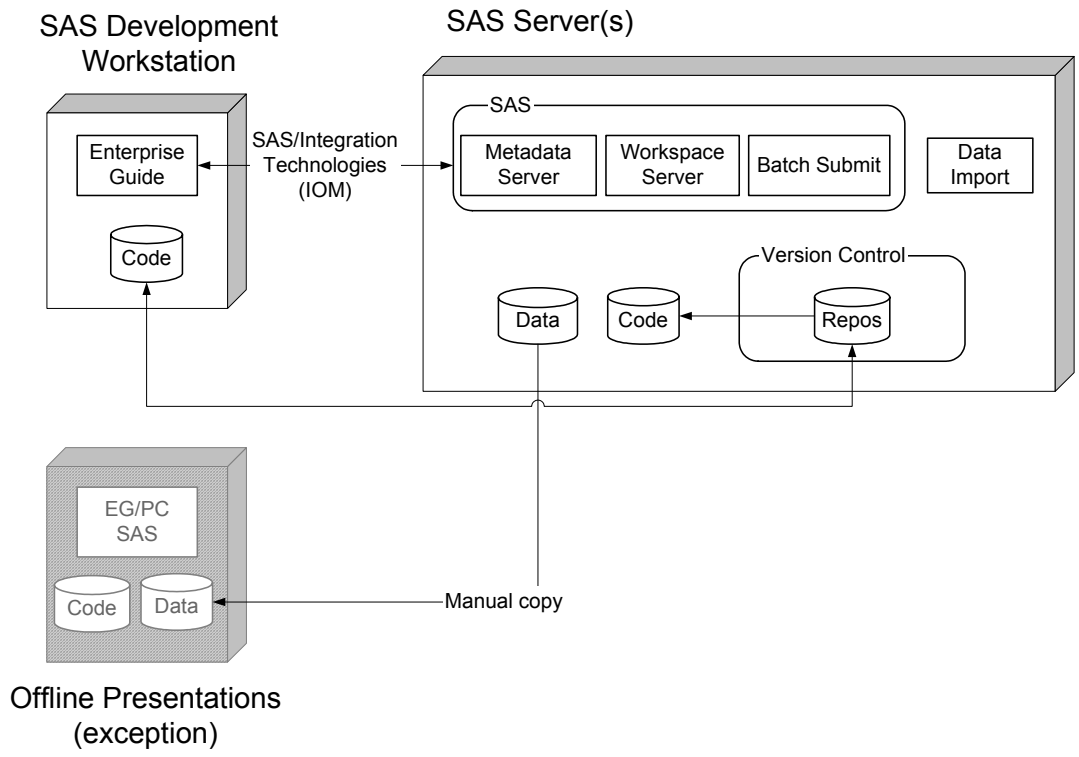

Figure 3. Hybrid Profile

In this diagram, typical SAS tool development relies on Enterprise Guide and server-based SAS (Workspace Server. This model supports the concept of centralized server resources. It should be noted that some use of batch SAS may also be the most effective way to generate production results; this is an area for additional analysis.

There is a certain amount of complexity involved with this proposed architecture, so the above diagram only shows variations that are supported by high priority business needs.

One option that clearly deserves analysis and consideration is to apply the new SAS9 multi-tiered architecture. The SAS Institute has made a large investment in these new products and technologies, and these are the future of SAS for some time to come. It must be recognized that some of the new products - such as Information Map Studio, Web Report Studio, and the OLAP server - are aimed at the general Business Intelligence problem area, and are not necessarily appropriate to the development of tables, figures, and listings. However, there are aspects of SAS9 that can be used.

In terms of study development, the logical architecture would be to use Enterprise Guide (EG) as the main tool for Statisticians and Statistical Programmers, and to use the new SAS9 servers for server-based analytics. Enterprise Guide is recommended as the client because it is able to seamlessly make use of a remote SAS server. It also offers some useful wizards, and is openly extensible using SAS-provided interfaces. Finally, Enterprise Guide is the strategic client tool in which SAS Institute is making the most R&D investment.

There are a number of benefits in using the new SAS multi-tiered architecture:

- Delivers immediate productivity benefits to the SAS users, particularly those that are currently working with SAS on Unix.
- Makes sensible use of both the desktop and the server. The desktop is used for code editing, reviewing results, and data exchange. The server is used for SAS computations, which are the most CPU and disk intensive operations.
- Provides excellent support for centralization of SAS, consistent with both the business vision of a global system and the overall IT consolidation strategy.
- SAS Management Console allows management of license information, installation/ configuration information, which servers and applications are out there, and some ability to monitor SAS servers.
- Keeps Covance in the main stream of SAS usage, and projects an image of being technologically current to Covance customers and other parties.

Further, this option may create an infrastructure that is useful for other purposes. For example, some ad hoc query needs could be addressed through clients such as Enterprise Guide and the add-in for Microsoft Office. These clients could validate against the Metadata Server to access summary information or study data that is housed in study libraries. Data reviewers could access STUDY123 and see any number of datasets and use these clients to query the data. Each study would have to have a unique libname associated with that study.

CURRENT STATE OF THE MIGRATION

As of the writing of this paper, we are about to embark upon a 10-week benchmark study (which will be nearly complete at the time of the presentation). SAS 9 presents opportunities for Covance to think about how its business units employ SAS software, and it is enough of a difference to warrant testing and proof that the

architecture will not only support the needs of the organization and be easy to manage from an IT support perspective, but also advance the vision for creating a global platform for Covance.

The purpose of the benchmark study evaluation is to mitigate risks that might be associated with the proposed architecture; to clarify what technology needs to be built versus defined as a process (or processes) and to provide metrics that can be extrapolated as a basis for estimating return on investment (ROI).

The benchmark study will seek to answer a number of questions about how SAS 9 and the metadata server can help with a number of business related processes. The following areas are candidates for additional automation through development of custom integration or tools to support statistical studies:

- **Study management –** Creation of a new study area and authorization
- **Version control –** Integration of functionality within Enterprise Guide; use of version control, file management practices
- **Output management –** Tools for producing draft/final output; Use of more SAS capabilities for document formatting, more efficient assembly of output, file management practices, randomization
- **Import management –** Tools to standardize procedures around bringing data into a study; Import of other study files, use of CDISC SDTM data model
- **Metadata management –** Tool support for reuse of study metadata

CONCLUSION

As is the case with any technology advancement, organizations should take a critical look at whether or not an improvement in business processes should accompany the enhancements made viable by technology. Covance took the migration of SAS from Version 8.2 to SAS 9 as just that opportunity. Through careful evaluation of the technology in the context of the business, and by conducting an assessment to surface how these changes could improve the way they worked and delivered results to their clients, Covance architected a solution that would not only satisfy the immediate needs of the organization, but prepared them for a future where change is inevitable. It should be noted that the deliberate planning, phased execution and testing helped Covance move the vision forward while maintaining the integrity and robustness of its operational capability in a highly regulated environment.

ACKNOWLEDGMENTS

The authors would like to sincerely thank Tim Walters and Jeff Wright for their guidance and thoughtful review of this manuscript.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Please feel free to contact the authors at:

Greg Nelson, President and CEO
greg@thotwave.com
ThotWave Technologies, LLC.
2054 Kildaire Farm Rd, #322
Cary, NC 27511
800.584-2819 – Phone/ Fax

Michael Walega, Executive Director, Global Biometrics Operations, Late Stage Development Services
Michael.Walega@covance.com
Covance, Inc.
210 Carnegie Center
Princeton, NJ 08540
609-452-4150 (voice)
609-520-1754 (fax)