A Case Study in the Use of SAS/PH-Clinical® for an Electronic Regulatory Submission of a Megatrial

Kathy Boussina, Dori Emmrich, Guy Pawson, and Alan Hopkins
Genentech, Inc.
South San Francisco, CA

Abstract

SAS/PH-Clinical was used for a Product License Amendment (PLA) of the GUSTO (1,2) trial, an international megatrial comparing four thrombolytic strategies in the treatment of acute myocardial infarction. Using commercial software, Genentech was able to prepare an electronic regulatory submission within 8 weeks of receipt of the final database. The overall package included text, graphics, programs, data sets, supportive documentation and usage instructions contained on two 550 megabyte CDs. SAS/PH-Clinical allowed both medical and statistical reviewers to review and analyze the data.

Introduction

The Global Utilization of Streptokinase and Tissue Plasminogen Activator for Occluded Arteries (GUSTO) trial was conducted under an investigator-sponsored IND to evaluate survival benefit among four thrombolytic (clot-dissolving) strategies. With 30-day mortality as the primary endpoint, 41,021 patients were randomized to one of the four treatment groups. Patients were recruited from 15 countries and 1081 hospitals. The trial used an experimental accelerated dosing regimen of Activase® that had been piloted in other studies (3-5). A subset of 2431 patients were also randomized to an angiographic substudy comparing coronary artery patency and mortality rates. Duke University served as the data and statistical coordinating center. At the end of the study, Duke University provided the data to Genentech. The data were not available in-house until the study was unblinded.

In order to provide the data to the FDA within a reasonable timeframe, a significant amount of planning and coordination occurred prior to receipt of the final database and even prior to the announcement of the study results. The last patient in the trial was randomized on February 22, 1993. The public disclosure of the GUSTO results took place on April 30, 1993. The results demonstrated a statistically significant 30-day mortality benefit with the new accelerated dosing regimen of Activase® over the three other thrombolytic strategies. Since this was an experimental, unapproved dosing regimen, a Product Licensing Application (PLA) supplement to modify the labeling was required. On August 13, 1993, Genentech demonstrated the
SAS/PH-Clinical software to the FDA using the GUSTO angiography substudy data and it was decided to use the software as a "pilot evaluation project". We received a finalized database from the GUSTO Coordinating Center on August 17, 1993. The submission was completed 54 days later. Close coordination with the FDA Center for Biologics Evaluation and Research (CBER) was essential to this process.

Goals and Objectives

The primary goal of this project was to provide a timely turnaround in the form of a PLA. We not only wanted to provide the FDA with the submission within a short period of time from receipt of the database but also hoped that the tools included would assist and thereby expedite the review.

In order to achieve this goal the following objectives needed to be met:

- Provide data review capabilities to medical and statistical reviewers using commercial software
- Provide submission documents and data in an electronic format

Materials and Methods

To achieve the goals and objectives outlined for this submission, a multidisciplinary team was formed at Genentech. This team consisted of individuals from the following disciplines: biostatistics, programming, data management, clinical research, regulatory affairs, document processing, and computer information systems. At the FDA, personnel from the Center for Biologics Evaluation and Research (CBER) divisions of scientific and management information systems, biostatistics and epidemiology, and application review and policy reviewed the CAPLA proposal. SAS Institute provided support through marketing, development, and SAS Consulting Inc.

Providing a Data Review Tool

Hardware and Software Selection

Several factors were considered in the selection of hardware and software.

The choice of software was driven by two primary considerations:

1. The desire to use commercial software, and
2. The need to provide a flexible analysis tool suitable for both medical and statistical reviewers.

Commercial software would allow us to take advantage of the training and documentation provided with such products. We hoped to obtain in-house training that would then prepare us to train the individual reviewers. In addition, since Genentech is located on the West
Coast, we had concerns about our ability to supply technical support for review software used in this submission. We needed a stable, fully tested, reliable tool. We selected SAS/PH-Clinical because of SAS Institute's long-standing reputation in the industry and because of the software's strong analysis capabilities. Using SAS/PH-Clinical took advantage of our own in-house SAS software expertise which proved useful in the development of our application. We were able to provide training to individual reviewers along with both SAS software manuals and documentation specific to our data. The proximity of SAS Consulting in the Washington, DC, area was another advantage in terms of additional technical support. CBER consented to use SAS/PH-Clinical software as a "pilot evaluation" project.

The basic functionality built into SAS/PH-Clinical is the capability to browse patient data, create patient listings, create patient subsets, derive new variables, create data tabulations, perform exploratory statistical analyses and view, execute, or copy Genentech-supplied programs and view the associated tables, graphs, and listings. The ability to view patient level data and to tabulate and summarize data seemed well suited for both medical and statistical reviewers. The product is menu-driven and therefore does not require programming expertise. It does, however, have built-in access to SAS/Insight® and SAS/Display Manager® which might prove useful to someone with a statistical or programming background.

The CBER medical and statistical reviewers were using MS-DOS®-based software before this submission and were in the process of upgrading their hardware and software environment. We wanted to provide software that could eventually run on their in-house machines. The OS/2® operating system was recommended for this project by SAS Institute because of its superior performance compared with MS-DOS or Windows 3.1® platforms for large data sets. Individuals from SAS Institute spent a great deal of time and effort evaluating SAS/PH-Clinical performance using a preliminary GUSTO data set. OS/2 was selected after running several benchmark tests comparing DOS/Windows and OS/2 (see Table 1 below).
Table 1
Selected Performance Benchmarks Performed by SAS Institute for the GUSTO data (N = 41,021, Variables = 315)
(times are in minutes:seconds)

<table>
<thead>
<tr>
<th>Benchmark Test</th>
<th>OS/2 2.1</th>
<th>Windows 3.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Build Patient Group</td>
<td>11:02</td>
<td>24:01</td>
</tr>
<tr>
<td>(Subset Patients)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two-Way Table</td>
<td>1:12</td>
<td>1:47</td>
</tr>
<tr>
<td>Patient Listing</td>
<td>15:59</td>
<td>29:21</td>
</tr>
</tbody>
</table>

SAS Institute also assisted us with specifications for hardware. High-end PCs were chosen for this project because of our performance requirements. We selected a 486 microprocessor-based system with a 66 mHz clock rate and 32 megabytes of DRAM. In addition, the machines were configured with dual 1 gigabyte hard disks with SCSI interface, a CD reader, and an 8 mm tape backup system. This configuration provided maximum performance and storage for large data sets.

Documents

It was agreed to submit all documents in WordPerfect format since this is the word processor familiar to the individual reviewers at CBER. Included in WordPerfect were scanned images of all references (i.e., publications), scanned annotated Case Report Forms (CRFs), in-text tables and graphs, the final reports and a table of contents. The detailed table of contents served the dual purpose of navigation through the paper and electronic documentation.

Data

The data set for the parent GUSTO study contained 315 variables and was 72 megabytes in size. The angiographic substudy consisted of 2431 patients, 243 variables and was 7.2 megabytes in size. All data could be accessed through SAS/PH-Clinical. In SAS/PH-Clinical a data set is referred to as a "Patient Group." In addition to these two primary patient groups, we provided a total of 13 subset patient groups.
groups for ease of review and for performance considerations. The various patients groups consisted of selected variables on all records or a subset of records. Access to these patient groups was defined as "read-only" within SAS/PH-Clinical by setting up appropriate security groups.

Custom SAS software analysis programs were also supplied. A total of 157 programs consisting of 55,000 lines of SAS software code were supplied within SAS/PH-Clinical's program library component. These programs can be executed to reproduce the tables and figures that were part of the submission. The program name was tagged to the table name in the Table of Contents of the study report. In addition, a number of report table generating programs were modified to run using the current patient group in SAS/PH-Clinical. Thus, a medical or statistical reviewer could reproduce a table using a desired subset of patients. All tables and figures contained in the submission were available for viewing within SAS/PH-Clinical's Output and Graphics Libraries.

Performance Considerations

Because of the large amount of data contained in the GUSTO data set, performance was an issue during the analysis. In some cases processing times can be longer within SAS/PH-Clinical than when using regular SAS software code as a result of extra data checking code in the system. SAS/Screen Control Language (SCL) is used throughout and also contributes to the performance deficit experienced with SAS/PH-Clinical. The size of the GUSTO data set brought performance issues to the forefront enough to warrant revisions in the maintenance release that followed this submission.

SAS Institute incorporated the functionality to make optional the saving and restoring of a data set when accessing the library. For the GUSTO data set this takes almost 5 minutes.
Table 2 illustrates the performance issues encountered with the GUSTO data using SAS/PH-Clinical.

Table 2
Estimated Times for Reviewing Entire GUSTO data set

<table>
<thead>
<tr>
<th>Activity</th>
<th>Appx. Time (in minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retrieve entire Patient Group</td>
<td>5.0</td>
</tr>
<tr>
<td>Browse patient data with Drilldown</td>
<td>3.0</td>
</tr>
<tr>
<td>Go To Last Record when browsing data</td>
<td>0.5</td>
</tr>
<tr>
<td>2-Way Table</td>
<td>1.0</td>
</tr>
<tr>
<td>Ad Hoc Table (uses PROC tabulate)</td>
<td>60.0</td>
</tr>
<tr>
<td>Create subset patient group</td>
<td>10.0</td>
</tr>
</tbody>
</table>

Within SAS/PH-Clinical we strongly recommend that subsets of data be used whenever possible in order to reduce the processing time. In addition to those patient group subsets provided by Genentech, FDA reviewers were able to create desired subsets of variables or subsets of patients meeting certain criteria for a particular analysis. For the overall analyses where use of the entire data set was necessary, expected processing times were provided to the FDA reviewers for the more commonly used functions.

Submission Provided on Compact Disks

For archival purposes compact disk was the submission media requested by CBER. Genentech acquired the appropriate software, hardware, and in-house expertise to write the CDs for this submission. All WordPerfect documents were contained on one CD, and data sets, catalogues and libraries for use with SAS/PH-Clinical were contained on a second CD. All SAS software was also provided on CD. At the time of software installation, the data CD was copied to the FDA's hard drives for optimal performance. The documents continue to be accessed directly from the CD.

Quality Assurance

Several measures were taken to ensure the quality of the final submission. All programs produced final customized output that was included in the submitted report. The tables were never retyped and thereby the risk of error was reduced. The program output was reviewed and validated for consistency and results were compared back to prior analyses produced by the Coordinating Center and against SAS software database queries.
All programs were written and initially run on a UNIX server. Data and programs were then downloaded to the PC and all custom programs were rerun within SAS/PH-Clinical. The output generated on the PC was compared with the original UNIX server generated output using the UNIX "diff" command. This process not only verified the programs but also the data sets. Additionally, several patients were arbitrarily selected and a hard copy of the custom screens was compared with listings generated from the SAS software data sets. All components of SAS/PH-Clinical were tested by using the data from the final CD. Documents were viewed directly from the CD and compared with the originals.

Documentation and Training

We were able to take advantage of SAS Institute's on-site training program which allowed us to quickly become proficient at using SAS/PH-Clinical. Additionally, we were then able to provide training directly to medical and statistical reviewers at the FDA. We preferred this since we could provide complete orientation not only to SAS/PH-Clinical but to the organization and structure of our submission.

Genentech personnel trained individuals from both CBER and CDER in the use of the data review software. Approximately ten FDA medical and statistical reviewers were trained over a three-day period that consisted of hands-on sessions lasting approximately two hours. Multiple sessions were held with primary reviewers. During training, reviewers were familiarized with the organization of the submission package, the retrieval of documents from the CD, and the organization of the data and programs within SAS/PH-Clinical. A training manual was provided to facilitate the process. This manual contained excerpts from the original submission document that were considered especially helpful to the electronic component. An alphabetical list of all variables with a brief description and explanation of variable coding was included among the modules that composed this manual along with an annotated directory tree and annotated CRFs.

SAS Consulting was contracted to provide technical support to the FDA for SAS/PH-Clinical. This support includes response to the FDA personnel's questions regarding software, defining new devices and new users, and installing maintenance releases. The institute would also provide any additional SAS/PH-Clinical training necessary. Any data-related questions were handled by Genentech.

SAS Institute representatives were present during the training sessions and provided technical feedback. The training sessions allowed for a unique opportunity to detail the structure of the data and organization of the submission.

531
Discussion

Genentech was able to provide an electronic regulatory submission of a megatrial to CBER within 60 days from receipt of the final database from the GUSTO Coordinating Center. During that same time period, hardware and software were configured and training was provided to the FDA reviewers. Leveraging commercial software allowed us to focus on the content, organization, and quality of the submission and the accompanying documentation and usage instructions instead of on systems development.

The data review software would ideally be used from study initiation and throughout the course of the clinical trial. SAS/PH-Clinical was used as a back-end for our submission process, because Genentech did not receive the final data until the study was unblinded. This prevented us from taking full advantage of feedback from users prior to delivery of SAS/PH-Clinical with the GUSTO data to CBER. This step would have helped to identify in advance issues specific to the submission.

By providing the FDA with SAS/PH-Clinical, medical and statistical reviewers had the tools needed to replicate our analyses, generate new tabulations and analyses, generate custom data listings, and view individual case data. Early and continuous communication with CBER representatives and reviewers made this effort possible and contributed to the success of the software and submission installation. Performance considerations inherent in the current software may dissuade the reviewer from taking full advantage of SAS/PH-Clinical's functionality when accessing large data sets, however. Specifically, the ad-hoc tables components within SAS/PH-Clinical was found to be useless for the GUSTO data because of slow processing time. We look forward to discussions with CBER representatives once the review is complete to evaluate the process and the submission.

We must note that the design of this study was very simple, a single drug administration with in-hospital and 30-day follow-up. To date, we have no experience using SAS/PH-Clinical for a more complicated study design or for the submission of multiple studies.

Recommendations

Some software recommendations that would enhance this and other submissions are outlined below.

1. Improve batch processing capability

The portability of SAS software allowed programs to be prototyped on a UNIX server running SAS System Version 6.07 and downloaded to the PC. This was necessary as several
programmer/analysts and statisticians worked on this project. We had hoped to do the complete SAS/PH-Clinical development on the UNIX server but did not because a production version for that platform was not available at the time. This scenario required that all downloaded programs be rerun within SAS/PH-Clinical on the PC, a process that could not be achieved through batch processing and proved quite time consuming. A client/server environment would have alleviated some of these issues. This architecture would also be useful in situations where world-wide drug development requires distributed data access.

2. Provide electronic linkage of SAS/PH-Clinical data, documents, and CRF Images

This submission provided two environments for the reviewer: SAS software for analysis and WordPerfect for documents. Ideally, these would be within the same environment or the two would be "linked" through software. In addition, document imaging, i.e., scanned images of individual CRFs, should be part of future submissions and optimally accessible within the SAS/PH-Clinical environment through drilldown. This feature is not currently available in the SAS system, but is essential to a complete electronic submission.

3. Improve documentation with SAS/PH-Clinical

We believe the capability to provide more user-supplied documentation should be available within SAS/PH-Clinical. This could be achieved through customized help screens. The training binder that we supplied the FDA for assisting with the review would preferably be available through the customized help screens. Additionally, output generated from SAS/PH-Clinical needs to be more thoroughly documented. For instance, default titles identifying the patient group used and any subsetting performed would be useful.

4. Improve performance

Within SAS/PH-Clinical performance needs significant improvement for use with large data sets. Some performance issues have already been addressed since the initiation of the GUSTO submission and were instigated by use of this large data set. SAS Institute incorporated some modifications with the first maintenance release. There are still some areas where streamlining would be useful, such as the elimination of code to check for duplicate records prior to every procedure. This check occurs repeatedly for the same patient group and is particularly unnecessary for patient-level data. Similar repeat data step processing occurs for the identification of unique values when using subsetting criteria. With large data sets these steps can be costly in terms of performance.
5. Provide library structure capability

The ability to save SAS/PH-Clinical generated output and the associated SAS software program code into the library components is a strength of this software. The link, however, requires that the user supply consistent names to both the output and the program. A more automated, less error-prone linkage would be advantageous. Furthermore, the library components do not allow for any type of directory structure which makes managing large numbers of programs quite difficult. Objects within the library are (by default) sorted according to date and time stamps, but with 157 programs it was essential to sort them alphabetically. Outside of SAS/PH-Clinical we were able to sort the directory accordingly. Since all programs were stored within the same directory, we developed a naming convention early on to help in the organization of programs. More options in structuring and organizing the libraries are needed especially when considering that most submissions incorporate a number of studies.

6. Standardize user interface

The user interface was also clumsy at times. Scrolling and windowing are not consistent with other software applications, e.g., Windows. Additionally, the mechanism for building custom screens requires SAS/FSP software which is error prone and tedious. A follow-up visit with SAS Institute identified many of these issues as ones that will be addressed with SAS System Release 6.10. For this release, SAS/PH-Clinical will be redesigned using Frame and Desktop technology. Windowing will conform to windowing standards. Customized windows will be more readily available allowing for customized applications to be accessible from the main menu. The look and feel of the product is expected to be enhanced significantly. We look forward to these future releases and to using SAS/PH-Clinical more widely in our clinical trials data processing environment.

Acknowledgments

The following persons contributed to the success of this project: Art DeVault, Ph.D. Allene Dodge, Fong Wang-Clow, Ph.D., Gracie Lieberman, Gareth Creasey, Jan Van Gelder, Judy Stein, Daisy Wong, Nancy Milliken, Ted Love, M.D., Kerry Lee, Ph.D., Lynn Woodlief, Gary Klonicki, Sandra Schlotzhauer

References


