SAg. APPLICATIONS IN THE HEALTHCARE INDUSTRY

J. Jon Veloski, Jefferson Medical College of Thomas Jefferson University
William J. McDonald, Blue Cross & Blue Shield of Delaware, Inc.

ABSTRACT

A review of the proceedings of the past SUGIs reveals steady growth in the use of SAS software throughout the healthcare industry. This paper summarizes applications of SAS software to collect and manage patient medical data, drug safety data, insurance claims data, and other healthcare data. Also highlighted are ways in which some of the unique capabilities of SAS software have been used creatively to address the broader issues of quality control in medical care, financial analysis, statistical analysis, and healthcare policy analysis. Current applications under development and plans for the future are described.

Six important factors might explain why SAS software has been and will continue to be an indispensable tool for the development of innovative information systems in the healthcare industry:

1. The changes that arise from continuous growth in medical knowledge and technology demand flexible information systems that have the shortest possible development times and at the same time can be modified easily.

2. Increasing pressures to reduce costs often imply a need to minimize development effort.

3. Competition among providers, suppliers and insurers is creating a need for the most creative and effective information systems.

4. Standards of quality in healthcare information systems have always been high and cannot be compromised in order to reduce costs or save time.

5. Regulation by federal, state and local authorities continues to challenge many organizations' capabilities to deliver timely and appropriate information at minimal costs.

6. Both the volume of data and diversity of users within many healthcare organizations create a need for powerful, multi-purpose software.

In order to provide an overview of representative applications of SAS software in this industry we reviewed the published proceedings of the meetings of the SAS Users' Group International (SUGI) since 1983. What follows is a summary of selected papers drawn from these proceedings, as well as several current applications that we have been involved in, which illustrate a few of the ways in which SAS software can be used to tackle challenges in the industry.

Drug Safety Data/Clinical Trials

As evident in the list of references provided at the end of this paper, there have been substantial uses of SAS software for data reporting/analysis as well as data management in pharmaceutical firms. There, SAS software is regarded as the standard. Adequate summaries of many of these applications are available in SUGI proceedings, therefore we will mention a few that illustrate how SAS software can be used.

Tappe (SUGI 10) described some the typical problems that are to be encountered when data must be pooled across multiple drug studies in order to gain statistical power by which to evaluate drug efficacy and safety. Of particular value in this paper is the outline of a data structure that can be used to organize the raw data that support studies of drug safety. Bingham-Barnes (SUGI 10) provide further detail on the management and analysis of multiple, large scale clinical trials. Finally, Rosenberg (SUGI 13) provides a useful overview of clinical applications in pharmaceutical firms.

SAS software is also used widely to manage clinical trials in other settings. A major challenge addressed by Branagh (SUGI 13) is the need to minimize the cost of
writing SAS code for the routine editing of data sets from very similar clinical trials. He describes the very logical concept of a SAS code generator (SCOGEN) which reads a flat file containing the specifications of the variables under study and writes simple, syntactically correct SAS code. The flat file for each clinical trial can be created by secretarial staff who key in the following parameters for each variable in a particular study:

- SAS variable name
- LABEL
- INPUT location
- LENGTH
- INFORMAT
- FORMAT
- Acceptable values

One can further imagine how such routine data could be accumulated in a master dictionary for all studies so that a program could select the parameters for relevant variables from the master dictionary to write the custom code for later studies. Further developments in SAS code generators are also reported in the proceedings of SUGI 14.

The opportunity to use PC/SAS® as a front end to collect and manage clinical trial data was described by Johnson (SUGI 12) and Hope (SUGI 13). By using WINDOW and DISPLAY on a PC it becomes possible to edit and manage data in a “user friendly” manner that was almost unimaginable in the mainframe environment. The ability to create DBF or DIF files, and of course the RLINK procedure all assure that the clean data can be expeditiously moved to virtually any other computing environment as need be.

**Claims and Membership Files - Blue Cross/Blue Shield**

SAS software enables decision-makers with urgent needs and diverse technical capabilities to exploit valuable data that are already being automated routinely in an organization. It is not uncommon to find capable and creative managers frustrated as they wait for reports to be coded by central programming staff who are overwhelmed by requests from all parts of an organization. Faced with a rapidly growing queue of user requests for reports and analyses, Blue Cross & Blue Shield of Delaware chose SAS software in 1983 as the primary means of making claims and membership data available to a wide base of decision-makers. They established an information center to provide user training and consultative support, and more recently the historical claims data were loaded into System 2000®. User-written SAS programs are now used routinely, for example, to generate actuarial reports, provider profiles aimed at documenting billing patterns, account-specific utilization rates, demographic studies for marketing, market share studies and a variety of forecasts for corporate planning. The use of SAS software is so pervasive that, in 1988, the computer resources utilized by end-user computing began to approach those used by routine claims processing.

**Cost Data - Large Medicaid Claims Files**

Fox and Ray (SUGI 13) set out to find the most efficient way using SAS software to extract a complete claims history (approximately 15 million claims) for a subset (n=65,000) of medicaid patients having the diagnosis of diabetes mellitus. The ultimate goal of the study was to evaluate rates of hospitalization, emergency room visits and physician visits for these patients. The source was a multi-year claims file of about 50 million records. They evaluated the following four approaches: binary search, hashing, sort/merge and multiple FORMAT tables, and concluded by recommending a creative use of the SAS FORMAT in order to minimize processing costs. First, the patient identifiers of any claims with the relevant diagnostic codes were extracted and stored in a SAS FORMAT. Actually, multiple FORMATS were used in order to reduce the REGION size needed to create each FORMAT. Then, during a second pass against the claims file each of the 50 million claims was examined against the FORMATS to find out whether it was to be included in the subset of claims. Fox and Ray present a thorough discussion of the tradeoffs involved in the four alternative methods that might have been used to conduct this study, and illustrate the effective use of SAS FORMATS.

**Editing and Screening**

**Large Healthcare Data Sets**

SAS software is particularly well-suited to the retrospective analysis of large healthcare datasets when standards of statistical sampling are rigorous, screening criteria are complex and objectives are frequently changed. McDonald and Veloski (SUGI 13) described an approach that can be used to review data in the industry's
standard format, UNIBILL or UB-82, for inpatient and outpatient claims. Flexible edits were coded in either Boolean logic, or in the case of ICD-9-CM codes, using FORMAT libraries. Patterns of continuity of healthcare across multiple claims were evaluated with RETAIN statements and SET statement pointers. Statistical sampling was operationalized with the UNIFORM function or by DATA step coding when interval sampling with residual counts was needed. Used for review of Medicare, medicaid, HMO and outpatient surgery claims, the selection of SAS software ultimately assured that the data could be made available to many levels of users after the samples were drawn.

**Product Definition in Hospitals**

A valuable yet sometimes unappreciated feature of SAS software is the capability for users to write their own PROCs in another computer language. These can be linked to SAS software to enable the user-written PROC to operate on SAS data libraries just as any other SAS PROC. An example within the healthcare industry is PROC STAGE that has been used by hospitals and government to analyze hospital discharge data sets. PROC STAGE implements disease staging, a medically-sound system of logic which classifies patients' medical problems according to levels of severity within any of about 400 disease categories by examining vectors of ICD-9-CM diagnostic and procedure codes.

As can be imagined, PROC STAGE has been used to address healthcare issues such as within DRG analysis, hospital case-mix reports, provider mortality rate studies, resource utilization and quality assurance. One particularly innovative use is in the area of product definition in which hospitals are seeking a clearer understanding of the impact of patient severity on the cost of treatment, as they are forced to make competitive pricing decisions in a changing environment. For example, the management of a large, urban teaching hospital (Hospital A) became concerned when they learned that a local HMO had negotiated a contract with another reputable hospital (Hospital B) to perform angioplasty of the coronary arteries on certain HMO subscribers. What was surprising was that Hospital B was over 1,000 miles away, yet its negotiated price for the angioplasty including travel and accommodations for both patient and spouse was several thousands of dollars less that Hospital A's estimate of their cost for the procedure! Hospital A needed information in order to understand this incongruity. Might they be in danger of losing their competitive advantage in this procedure?

Disease staging provided a model for solving this puzzle and PROC STAGE was used to perform the analysis. A retrospective analysis of the charges for the angioplasties performed at Hospital A showed that the mean total charges had been $14,855, and the hospital planned to set their price slightly above this. However, analysis by disease stage indicated that the mean total charges ranged from $12,155 for patients of low severity, $15,063 for patients of moderate severity and finally $20,334 for very sick patients. The answer as to why Hospital B's price was set so low was actually simple. Any HMO patient well enough to await elective surgery at Hospital B, plan the air trip to the distant city and undergo the stress of travel would be unlikely to have a complicated illness, and therefore would probably cost less to treat. Hospital B could safely assume that any seriously ill patient, would be referred to a hospital nearer their home.

**Summary**

We have described several exciting applications in which SAS software is being used to address current challenges in the healthcare industry. We identified many valuable papers in the previous SUGI proceedings, and therefore encourage interested readers to review the enclosed list of references to locate SUGI papers related to their interest.

**For Further Information Contact:**

Jon Veloski  
Center for Research in  
Medical Education and Health Care  
Jefferson Medical College  
Thomas Jefferson University  
1025 Walnut Street  
Philadelphia, PA 19107  
USA  
Tel: 215-928-7967  
FAX: 215-928-5044

Note: *SAS, PC/SAS and System 2000 are the registered trademarks of SAS Institute, Inc., Cary, North Carolina, USA.*
SELECTED REFERENCES

PHARMACEUTICALS - DATA MANAGEMENT


Bauer, N.G. Designing A Drug Information System for The FDA Based on the SAS System, SUGI 13, 88:568-573.


Corcoran, C.E. SAS and The IBM Dialog Manager at Miles Pharmaceuticals, SUGI 11, 86:587-592.

Fondy, E.M. & Lee, J.Y. SEQURAN-An Interactive Sequential Randomization System for Clinical Trials, SUGI 9, 84:775-780.

Frueh, P.A. GENISYS - A Generalized Inventory System for Clinical Trial Data, SUGI 9, 84:651-655.


Note: References refer to the published proceedings of respective SUGI meetings.


Wolf, C.L. & Guthrie, B. Computerized Clinical Data Bank Networks, SUGI 11, 86:324-328.

**PHARMACEUTICALS - REPORTING/ANALYSIS**


Bourn, K.L. PROC GLM - Rearrangement of Output, SUGI 9, 84:661-666.

Costa, B. TABLETS - A Table Generator, SUGI 12, 87:1194-1197.

Crowe, J.T., Jr. & Brizendine, S.B. Interactive SAS Program Generators for Analyzing and Reporting Medical Research, SUGI 10, 85:844-847.

Crowe, J.T., Jr. & Malicz, T. SAS Programs for Reporting the Results of Bioequivalence Trials, SUGI 9, 84:589-591.


Lee, J.Y. & White, J.M. %SURVCOMP - SAS %MACRO for Descriptive Survival Analysis, SUGI 10, 85:1051-1058.


O'Connor, K.K. Rapid Response to The FDA Questions on An NDA, SUGI 9, 84:771-774.


Park, D.R & Ruberg, S.J. Adverse Reaction Incidence Tables for Clinical Trials, SUGI 8, 83:624-629.


**INSURANCE**

Armstrong, M.A. Using SAS to Compute Standardized Rates by The Direct Method, SUGI 8, 83:690-695.


BIOSTATISTICS/EPIDEMIOLOGY


Lee, J.Y. & White, J.M. %Survcomp - SAS %Macro for Descriptive Survival Analysis, SUGI 10, 85:1051-1058.

Pickle, L.W. Screening for Potential Confounders in an Epidemiologic Analysis, SUGI 12, 87:935-938.

DATA MANAGEMENT - OTHER


Johnson, G.R. Using Window and Display in the SAS System for PCs for the Patient Registry of the VA Coronary Artery Surgery Study, SUGI 12, 87:793-797.


Pratter, F. & Leger, M. Living With the AHA: The Construction of a Multi-Year Hospital Information Data Base Using the AHA Annual Survey and SAS, SUGI 8, 83:458-461.


Van Grevenbof, P. et al. A Data Base for Reporting on Patient Flow in a Clinical Group Practice of Medicine, SUGI 8, 83:313-316.

OTHER HEALTH CARE


Corliss, J.W. Nutrient Data Base Systems as Educational Tools for Allied Health Sciences Professionals using the SAS/FSP Product, SUGI 9, 84:540-545.


Tong, C.P., et al. Cooperative SAS Usage: University of Toronto and Toronto General Hospital, SUGI 8, 83:555-556.

Tong, C.P., et al. How the SAS System is used by Division of Cardiovascular Surgery at Toronto General Hospital, SUGI 9, 84:581-582.