The SAS® System in Dental Clinical Trials.

P R Jones - Unilever Research Port Sunlight Laboratory

Abstract

Unilever Research is the science behind one of the world’s largest manufacturers of personal oral health care products. The testing of anti-caries toothpastes culminates in large scale clinical trials designed to demonstrate their effectiveness. SAS® was chosen as the software to administer, store and analyze the data from such a trial which involved over 4000 subjects for three years and over 2.5 million data points.

During the trial the system evolved from a basic version written in SAS® 5.18 to a fully fledged user friendly system using many of the features of SAS® 6.06 in the VMS® environment.

The paper will review how SCL has allowed the user interface to develop, and illustrate how to move SAS® applications from the arena of the computer specialist to the novice user. A number of examples will demonstrate the changes made possible as SAS® developed and will include showing elements of a system currently in use in one of our clinical trials.

SAS® has proved a very capable tool for the analysis of dental clinical trial data, and for a wide range of administrative tasks. However its wider application within such trials requires the addition of a number of new functions. The final discussion will address these future requirements.

1. Introduction

The SAS® system has been used widely by the pharmaceutical industry for the statistical analysis of data. Data are conventionally stored within a relational database system. To a lesser extent SAS®, has been used as a central data management and storage tool for the complete cycle of a large scale clinical trial.

Unilever Research has been conducting caries clinical trials for over 25 years. We began to use SAS® as the storage and data handling tool for clinical trials in 1988, replacing a Fortran data management system. At that time I would say that we were in the minority of companies using SAS® as a repository for clinical data.\(^{(1)}\)

So why did we select SAS®?

- Flexibility
- Speed of Development
- In-House Expertise
- Accepted within the industry
- Technical Back-up
I aim to present our view of the SAS® system, what we use it for, how the upgrade to SAS® 6.06 has improved our system, and our future needs and views.

2. The Setting.

Unilever conducts large scale caries clinical trials about every 2-3 years. Such trials often involve over 4000 subjects whose dental status is monitored for three years resulting in over 2.5 million data points. Typically these trials are conducted on adolescents who are 11/12 years old at the start of the trial.

A trial is carried out with the co-operation of schools, allowing us to follow and locate the subjects over the 3 years of the trial. The trial includes 4 examinations, a baseline examination at the start and 3 annual follow-up examinations. At each examination a complete dental assessment of all the available subjects is made.

The most recently published of such trials took place in Lanarkshire, Scotland, between 1988 and 1992. A total of over 4000 subjects took part in the trial which proved the clinical superiority of toothpastes containing sodium fluoride over sodium monofluorophosphate ones.

This was the first of our trials in which the SAS® system was used for the full cycle of the clinical trial, from the initial recruitment of the panellists to the analysis. The objectives of using SAS® were to

- Provide all the necessary options within the systems
- Provide a system that didn’t require specialist skills to operate.
- Provide a consistent and easy to use interface.
- Reduce the paper work for the Administration staff.
- Improve the quality of the collected data.
- Have software systems that were easy to maintain.

2.1 The Data Management

The data processing of the trial was divided into two applications, one dealing with the administration aspects and the other with the management of the clinical data. The systems were developed in SAS® 5.18 under the VMS® Operating System. The systems required the following elements of the SAS® system: Base (3), SAS/AF® (4) and SAS/FSP® (5).

Under SAS®518 a number of key features were not available

- SCL (6)
- Flexibility in AF screens, selection lists, extended tables etc
- Psedit screens could not have programs behind them.
- PMENUS
With these limitations a number of jobs now handled by SCL had to be performed by either using the Macro Language or operating system calls. (In this case VMS.)

I will use the start-up phase of the trial to illustrate some segments of the systems, their functions/capabilities and their development for future systems.

The flowchart (Fig 1) illustrates the typical data flows involved in the start-up phase of a caries trial.

Figure 1 Flowchart of Start-Up Phase
2.2 Data Model

To help in discussions of the systems a simplified model of the data storage is shown (Fig 2.) The model is based on the concept of relational databases and the data are stored in 3rd Normal Form.

As the model is simplified the Entities illustrated correspond to more than 4 datasets. The model represents the relationship between the files within the systems.

2.3 The Admin and Clin Systems

One of the most crucial phases of any caries trial is the start-up phase. This phase consists of: the collection of informed consents, the baseline examination and the allocation of products. These functions are handled by the Admin System. The Clinical (Clin) System deals exclusively with the management of the collected clinical data, which is input after product allocation.

I will focus on four distinct areas within these applications and discuss the methods used in a particular area, their upgrading and future needs.

2.3.1 Recruitment and Entry of Consent Forms

To recruit panelists onto the trial, schools were invited to participate in the study and those that accepted, supplied us with a listing of pupils in the targeted age range. Their names and addresses were then entered into a flat ASCII file by data prep staff. This flat file would then be loaded into a dataset for the production of the invitation letter, which included the parental consent form.

As completed consent forms were returned, the administration records were updated within the Admin dataset. This data was then validated. Data validation and quality control checks required that users had access to the ASCII files while using the admin system. The external file handling was therefore quite complex and required careful consideration.

Users don't want to know operating system details. They want friendly, simple and easy to use systems, and should be given such. 'Upto half the cost of operating a computer system is in collecting and preparing data for processing' and 'almost every case of computer fraud via computers in recent years has involved tampering with input or inserting bogus transactions through normal input channels'.

653
To avoid these pitfalls we needed a sophisticated interface for the user. SAS® 5.18 had little in the way of features for handling external files. Often calls to the operating system were the most efficient way of shielding the user. Handling SAS® datasets was somewhat easier. Listing 1 illustrates a typical method of reading in the names of SAS® datasets in a library. Typically these macros could then be displayed on the screen giving the user a selection from which they could then enter a number indicating their choice. One SCL command now achieves more than this.

**Listing 1**

```sas
proc contents data = datalib._all_ memtype=data out=dir noprint; run;
data dir1;
set dir;
 by memname;
 if first.memname;
 keep memname;
run;
%let dlist01= %let dlist05= %let dlist09= %let dlist13= %let dlist17=
run;
data _NULL_; set dir1 end = eof;
 length dstri $ 73;
 retain dstri ' , mn0 1 colno 0;
 i pos=colno*9+1;
 substr(dstri,i pos,9)=memname||' ';
 colno=colno+1;
 if mod(_n_,8) = 0 or eof then do;
 mname=compress('dlist'||put(mn0,z2.0));
 call symput(mname,dstri);
 mn0=mn0+1;
 colno=0;
 dstri=' ';
 return;
end;
run;
```

When SCL was introduced one of its primary strengths was to shield the user from external files. Part of the reasoning behind this was the introduction of MVA with release 6.03. The introduction of external file handling capabilities gave the end-user such features as selection lists, browsing and editing of external files within SAS®. Such a list of features has inevitably lead to higher expectations from users, as has the general introduction of systems such as Microsoft Windows in industry.

Most of the external file handling problems have been solved by the SCL functions present upto release 6.07. A few elements which would be especially useful in the clinical area are

- A set of audit trail functions for datasets (see later discussions)
- An improved environment for the editing of none-SAS® files.
2.3.2 Product Allocation

Product allocation in our caries clinical trials takes place before the entry of the CRF (clinical data) and the validation and Q.C Checking. The random allocation is however stratified by a number of factors which are collected at the baseline examination.

The original method for allocating products was paper based, with SAS® providing a random listing of product codes within strata. The interface in SAS® 5.18 was not considered robust enough to cope with this crucial stage, so the paper based system described was used. Each subject's stratum was calculated manually. After subjects had been assigned to a product on this listing, their product assignments were then entered into the admin system.

More recently (ie SAS® 6.06 onwards) the product allocation procedure has been completely automated. Fig 3 illustrates a typical input screen for product allocation. The screen includes verification checks, validation checks and the automatic assignment to the correct stratum leading to random product allocation.

![Product Allocation Screen](image)

**Figure 3 : Product Allocation Screen**

By adopting this new procedure we have reduced the mistakes that could occur, and speeded up the allocation procedure. This process can also now automatically cope with certain problems, such as siblings, i.e two participating children living in the same household must be assigned to the same toothpaste group. The end-user has a reduced amount of paper work and the process of quality checking is speeded up.
2.3.3 Input and Validation of Clinical Data

The input, verification and validation of clinical data has been described in a number of SAS® papers before, (7) and will not be dealt with in detail here. Our standard method was data verification through double-punching, with data validation then taking place. Normally the data was processed as ASCII files, with the checking being done in batch mode.

Historically this method had always proved to be a popular and reliable option, with low error rates. Nevertheless with the introduction of SCL behind Fsedit screens it was clear SAS® could be used for data Entry Screens. The possibility that data could be captured at source with cheap and portable PC’s became a real option. This however could not be implemented at the time because SAS® 6.06 was not available under DOS.

Instead a half way stage was implemented for a trial. This involved on the main-frame the setting up of Fsedit screens which were designed to closely mimic the layout of the CRF. The advantages from this were:

- Entry staff required less training, ie no need for special data prep staff.
- Immediate feedback on errors-reducing the cycle time.
- Visual checks may notice discrepancies.
- Data Entry staff prefer this type of input.

An Example of one screen from the input system is shown in FIG 4.

**Fig 4**

```
FSEDIT

Command ==> 

UPPER RIGHT
| 1.7 | 1.6 | 1.5 | 1.4 | 1.3 | 1.2 | 1.1 |
| domlb | domlb | domlb | domlb | domlb | domlb | domlb |

--- --- --- --- --- ---

UPPER LEFT
| 2.7 | 2.6 | 2.5 | 2.4 | 2.3 | 2.2 | 2.1 |
| domlb | domlb | domlb | domlb | domlb | domlb | domlb |

--- --- --- --- --- ---

Log : ___________
Log : ___________
Input : ___________
```

This screen allowed two input modes, field mode and line entry mode. In field mode the user entered the data directly over the fields by copying the CRF, while in line entry mode a command line is used to record the data, with a record of the last two entries being displayed. The data screen included full validation.
Remote Data Entry (i.e. data capture at source) is seen as the next logical step for improving the quality of captured data. It can offer the instant flagging of errors and longitudinal data validation. Apart from keyboard input, a number of other input methods might be used for the remote capture of data in a caries trial. However for this SAS® needs a more flexible interface for communicating with external devices, particularly for PC based systems.

As double data entry is still considered the best method for the verification of data, the introduction of a simple mechanism for double data entry with Psedit screens would be a worthwhile feature.

2.3.4 Menuing Systems

When a system is designed, a method for linking the sequence of jobs is required. This method is usually a menuing system.

In version 5.18 simple menus created an adequate user interface to the system. (Fig 5) With 6.06 application developers were given a whole new set of tools, one of which was the pushbutton. (Fig 6 illustrates a pushbutton menu.) With the movement within industry towards windowing systems (be it X-Windows or Microsoft) the mouse has become a common tool for the end-user.

The figures below illustrate the development.

![Main Menu](image)

**Figures 5 & 6:** Menu Screens
As Menus are often the first screen within a system, they must be friendly, consistent and easy to use, to encourage the use of the application. Our users prefer the ease and speed through which applications can be navigated with point and click menus.

As for the future, we already have 3 powerful tools for menu creation in SAS®. My suggestion is that it would now make more sense to consolidate the menuing functions within the menu entries in catalogs. One method for this would be the addition in the Gattr window of a single-station choice group allowing the selection of types block, text or pushbutton.

3. Discussion

The SAS® system has proved to be a successful tool for the management of caries clinical trial data within Unilever for the last 5 years. I can say this, because the systems have been regularly used for the last 5 years with no major problems being reported by users. As well as the areas detailed earlier, SAS® has been used for items including: adverse event reporting, CRF printing and letter printing.

It has offered a stable environment, while being progressive enough not to become obsolete. A similar system has now been ported over to a PC and can offer a full range of services while still being mobile. As an application developer the ability to quickly prototype a system has been invaluable in satisfying users needs.

The key features that made a difference to our users when considering SAS® 6.06 over SAS® 5.18 were

-Selection Lists
-PMENUMS
-A Consistent Environment on different Operating Systems
-Visual Appeal (,support for 'Windows')
-The ability to manipulate external files within the system.

From our view point the main areas of weakness for the SAS® system are:

In SAS® 6.07 security for datasets was re-introduced via a password system. This system does not deliver the required protection for Good Clinical Practice compliance on a PC based system, so system software needs to be added to augment the SAS® systems security procedures.

Remote data entry is seen as the future method for data collection. If the data is to be collected at source, a real time data collection module may be needed for the SAS® system. This could include links to interfaces such as the RS232, with tools for protocol and data management being available within the tool. It is extremely difficult to do real-time processing of data with the SAS® system. There is a large pre-clinical market that could be hooked into the SAS® systems power with the right interface software.
As noted earlier, auditing is a crucial aspect of a computerised drug system. CANDA's in particular will require that systems can audit the data trail. Without actually employing a relational database it not possible to quickly produce a system which can cope with roll-back mechanisms and automatic audit trailing.

With CANDA's around the corner (1995 is the target date[^9]), I hope that the SAS® Institute can co-operate in producing a standard for industry, and more importantly a regulatory authority excepted standard.

**4 Conclusions**

SAS® will continue to be used to as the management tool for the forthcoming trials. This is possible because the systems used produce an audit trail based paper records. However if we are to move to a paper-less system, including remote data entry the SAS® system will need an enhanced ability for auditing and control of data.

**5 References**

8. ADVANCS : An Automated Data Validation and Correction system, J Bondy et al, SUGI Proceedings.

**Trademarks**

SAS® software, SAS/AF® software, SAS/FSP® software and MVA® are all trademarks of SAS® Institute Inc.
VMS® is a trademark of the Digital Equipment Corporation.
Microsoft, MS and MS-DOS are all trademarks of the Microsoft Corporation.