

## SAS®-Based Data Management System for Clinical Drug Trials

Jean Bitney, A. H. Robins

Data collected during clinical trials at A. H. Robins Company are entered into a computer file and maintained in permanent SAS® data sets. A number of different types of profiles are routinely collected. Demographic, adverse experience, electrocardiogram, physical examination, clinical laboratory, and concomitant medication data profiles are similar for most clinical trials. Efficacy, dosing schedules, and vital signs profiles may vary within project, as well as across projects. Our data management system must be structured in a manner such that standards are easily maintained, where applicable, and still retain flexibility to handle project-specific profiles.

To fulfill these requirements in 1983, a CMS/SAS data management system was developed to define the SAS input record and create a SAS program to read this record. This system consisted of 3 CMS EXECs, 3 SAS skeletons, and 5 SAS program modules. One CMS EXEC module served as the system driver providing 6 options displayed in a menu. This driver module provided a conversational environment to elicit responses from the user, to validate the responses, and to pass control to another module. The SAS skeleton modules were edited with CMS commands, and the resulting files were executed as SAS jobs. A single selection from the menu consisted of multiple complex transfers from module to module. Application of the system revealed several shortcomings: system maintenance was difficult, enhancements to the system were delayed or avoided, and a high level of programming skill was required to accomplish these tasks.

After the installation of base SAS software version 82.3 at our company in the fall of 1983, the system was modified extensively. With the application of interactive techniques, SAS/FSP® software, and the macro facility, the CMS modules were replaced; the SAS skeleton programs were eliminated; and the number of required SAS programs was reduced. The modified SAS macro data management system consists of multiple SAS macros each related to a task. A help facility replaces the menu. Single or multiple tasks within a single project are executed during a session.

The modified system performs tasks in the following 4 general areas:

1. Profile definition.
2. Data dictionary table lookup and new variable definition.
3. Format code and description table lookup and new format definition assignment.

#### 4. Preparation of project programs and documentation.

The profile definition begins with the assignment of variable names to the data items collected. Variables are selected singularly or from a predefined group of related variables. A single profile contains from 1 to 70 variables. A profile defines one 80-column input record. A project has from 1 to 20 profiles. A SAS/FSP screen, Figure 1, is displayed for the user to select variable names. The record number is automatically defined as are the first 3 variable names, CARD, STUDY, and PATIENT. Additional variable names are added to the list in positions 4-15. Variables in positions 16-70 are added through Screen 2, Figure 2. Help information pertaining to this task is presented on Screen 3, Figure 3. Key variables, those employed to link to other profiles (data sets), are identified and ordered. In Figure 1 the 3 key variables are STUDY, PATIENT, and VISIT, indicated by 1, 2, and 3, respectively. The user instructions are presented on the screen. For example, in Figure 2 the user is instructed to press the PF 6 key to present HELP information; press the PF 10 key to continue. A permanent SAS data set of profile definitions is maintained for each project.

Once a series of profiles for a project are defined or the task is ended, the attributes of each variable are added with a table lookup facility employing a data dictionary. The attributes identifying each variable are record number; variable name, type, order, and label; length of input field; and input and output formats. Following a table lookup, the variable type and label, length of input field, and input and output formats are defined for variables found in the data dictionary. Variables not found in the data dictionary are missing these attributes. The attributes are added with the system definition task. The list of variable names and the associated attributes are stored in a permanent SAS data set for future reference.

The definition task allows the user to define the attributes of new variables. The SAS/FSP screen is shown in Figure 4. The fields for record number and name and variable order and name are protected fields and cannot be altered during the task. Two fields are provided for data entry notes pertaining to the variable. Input formats are defined for date and time fields only. Output formats are defined for variables as required. Output formats are subset in subsequent programming statements, and a table lookup employing a coding dictionary is performed to create a permanent SAS data set with project specific

format values and labels. Format values and labels are reviewed and modified through the screen shown in Figure 5.

A final task prepares project programs and documentation. A data entry and short encoding document provide a record for creating 80-column input records, Figure 6. Two SAS programs are generated. One SAS program reads the 80-column input record and creates SAS records to be added to a SAS data base. The other SAS program creates a permanent project specific format library.

Special features of the system facilitate the various tasks performed. For example, the REVISE task allows the user to modify a profile definition with the addition or deletion of a variable or a group of variables. The positions of variables following a revision are shifted accordingly. Another feature allows the user to select a group of variable names with identical root names with the entry of the root name preceded with an asterisk "\*." These variables are identified with programming statements and presented through a SAS/FSP screen to which the user identifies the number of items in the group. For example, the variable name \*PE selects the variable names PE1-PE14, Figure 1. Each variable in the series is defined with identical attributes following the data dictionary table lookup.

Because bioavailability projects at our facility are highly standardized, all documents for a single project are created with the task request DRUGM. A SAS/FSP screen is used to collect project-specific information for labels and sample time intervals. Then the SAS macro calls all tasks necessary to automatically generate the documents. The user is required to respond to the SAS/FSP screen once immediately after the system is called.

Two drivers call the system. A CMS EXEC driver establishes the interactive SAS environment, provides for a direct task call, and calls and executes the system SAS driver. The primary purpose of the CMS EXEC driver is for convenience. The SAS driver compiles the task files, calls the requested task or tasks, and displays the menu. Each task is stored as a separate series of macro statements. For ease of maintenance and revision, each macro is stored in a separate CMS file. The driver accepts from 0 to 5 task calls. Table 1 lists and describes the system calls. Examples of direct calls through the CMS driver are displayed in Table 2. The menu is presented with any of these calls:

```
DDCL (project) HELP
DDCL (project)
DDCL
```

A system SASLOG is not routinely saved in a CMS file. To save the SASLOG in a CMS file the LDG task is specified.

#### DDCL LOG (project) DEFINE

The SASLOG for the session is then saved in a CMS file named DRIVER SASLOG. Since all modules are SAS macros, the system may be driven with macro calls during any interactive SAS session.

The advantages of the modified system are:

1. The system code is task assigned and is easy to locate and revise.
2. Any experienced SAS programmer can easily maintain the system.
3. The enhancement of the system is accomplished through group assignments, quickly and easily.

SAS and SAS/FSP are registered trademarks of SAS Institute Inc., Cary, N.D., USA.

EDIT SAS DATA SET: MET072L1.List

```
COMMAND ==>>>          !SCREEN 1
                          !-----
                          ! OBS 1
```

#### SAS INPUT RECORD - VARIABLE NAMES

RECORD IDENTIFICATION NUMBER 02 NAME PEXAM

VARIABLE NAME	KEY VARIABLE ORDER
1. CARD	
2. STUDY	1
3. PATIENT	2
4. VISIT	3
5. PEDATE	-
6. MHSIG	-
7. WEIGHT	-
8. SBP	-
9. DBP	-
10. WEIGHT	-
11. PULSE 12. *PE 13. _____ 14. _____ 15. _____	

```
-----
PRESS PF 6 KEY TO ADD > 15 VARIABLES
PRESS PF 6 KEY TWICE FOR HELP INFORMATION
PRESS PF 8 KEY TO MOVE TO NEXT OBSERVATION
PRESS PF 2 KEY TO CONTINUE
```

FIG. 1. PROFILE DEFINITION - SAS/FSP SCREEN 1

EDIT SAS DATA SET

```
COMMAND ==>          !SCREEN 2
                   !-----
                   ! OBS 1
```

SAS INPUT RECORD - VARIABLES 16-70

```
16. ___ 17. ___ 18. ___ 19. ___ 20. ___
21. ___ 22. ___ 23. ___ 24. ___ 25. ___
26. ___ 27. ___ 28. ___ 29. ___ 30. ___
31. ___ 32. ___ 33. ___ 34. ___ 35. ___
36. ___ 37. ___ 38. ___ 39. ___ 40. ___
41. ___ 42. ___ 43. ___ 44. ___ 45. ___
46. ___ 47. ___ 48. ___ 49. ___ 50. ___
51. ___ 52. ___ 53. ___ 54. ___ 55. ___
56. ___ 57. ___ 58. ___ 59. ___ 60. ___
61. ___ 62. ___ 63. ___ 64. ___ 65. ___
66. ___ 67. ___ 68. ___ 69. ___ 70. ___
```

PRESS PF 10 KEY TO CONTINUE  
PRESS PF 6 KEY FOR HELP INFORMATION

FIG. 2. PROFILE DEFINITION - SAS/FSP SCREEN 2

EDIT SAS DATA SET:

```
COMMAND ==>          !SCREEN 3
                   !-----
                   ! OBS 1
```

SAS INPUT RECORD - HELP INFORMATION

- ARRAYED VARIABLES

TO ENTER VARIABLES THAT HAVE MULTIPLE ELEMENTS WITHIN THE SAME ROOT NAME AND THAT FOLLOW SEQUENTIALLY IN THE VARIABLE ORDER, PRECEDE THE ROOT NAME WITH AN ASTERISK (\*).

LIMITS - 1. ROOT NAME MUST BE 6 CHARACTERS OR LESS. THE SYSTEM WILL TRUNCATE ROOT NAMES THAT ARE 7 CHARACTERS.

2. THE RANGE FOR THE NUMBER OF ELEMENTS IS 2-99.

- KEY (BY) VARIABLES

KEY VARIABLES MUST FALL WITHIN THE FIRST 10 VARIABLES.

NUMBER THE KEY VARIABLES IN THE PROPER ORDER.

IF THE ASSOCIATED VARIABLE IS NOT A KEY VARIABLE, LEAVE THE BY VARIABLE IDENTIFIER BLANK.

PRESS PF 10 KEY TWICE TO RETURN

FIG. 3. PROFILE DEFINITION - SAS/FSP SCREEN 3

EDIT SAS DATA SET:

```
COMMAND ==>          !SCREEN 1
                   !-----
                   ! OBS 1
```

DEFINE SAS VARIABLE ATTRIBUTES

```
RECORD NAME _____
VARIABLE ORDER _____
VARIABLE NAME _____ LABEL _____
VARIABLE TYPE _____ NO. INPUT POSITIONS _____
(A,N,C)
INPUT FORMAT _____ OUTPUT FORMAT _____
DATA ENTRY NOTE _____
```

PRESS PF 8 KEY TO MOVE TO NEXT VARIABLE  
PRESS PF 2 KEY TO CONTINUE

FIG. 4. VARIABLE DEFINITION - SAS/FSP SCREEN 1

EDIT SAS DATA SET:

```
COMMAND ==>          !SCREEN 1
                   !-----
                   ! OBS 1
```

FORMAT CODES 1 - 13  
FORMAT CODES FOR: \_\_\_\_\_

VALUE	LABEL
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____

PRESS PF 6 KEY TO REVIEW FORMAT CODES 14 - 20.  
PRESS PF 8 KEY TO MOVE TO NEXT FORMAT CODE.  
PRESS PF 2 KEY TO CONTINUE

FIG. 5. FORMAT DEFINITION - SAF/FSP - SCREEN 1

CARD LAYOUTS

AHR-3070  
PROJECT; MET072  
02/04/85

CARD - 03 PEXAM

COLUMNS(S)	VARIABLE	A/N	LABEL STATEMENT	NOTES
1-2	CARD	N	CARD NUMBER	
3-6	STUDY	N	STUDY	
7-10	PATIENT	N	PATIENT	
11	VISIT	N	VISIT	DATA IS NOT AVAILABLE ON THE CASE REPORT FORM. PLEASE DO NOT KEY.
12-17	PEDATE	N	PHYSICAL EXAM DATE	
18	MHSIG	N	SIGNIFICANT PAST MEDICAL HISTORY 1 = YES 2 = NONE	
19-23	TEMP	N	TEMPERATURE - DEGREES FAHRENHEIT	
24-26	WEIGHT	N	WEIGHT (POUNDS)	
27-29	SBP	N	SYSTOLIC BLOOD PRESSURE	
30-32	DBP	N	DIASTOLIC BLOOD PRESSURE	
33-34	HEIGHT	N	HEIGHT (INCHES)	
35-37	PULSE	N	PULSE RATE	
38-39	RESP	N	RESPIRATION RATE	
40	PE1	N	PHYSICAL EXAM. - GENERAL APPEARANCE 1 = NORMAL 2 = ABNORMAL	
41	PE2	N	PHYSICAL EXAM. - HEAD 1 = NORMAL 2 = ABNORMAL	
42	PE3	N	PHYSICAL EXAM. - EYES 1 = NORMAL 2 = ABNORMAL	
43	PE4	N	PHYSICAL EXAM. - ENT 1 = NDRMAL 2 = ABNORMAL	

FIG. 6. DATA ENTRY AND SHORT ENCODING FORM

**TABLE 1 TASK CALLS AND DEFINITIONS**

TASK CALL	TASK DEFINITION
<HELP>	PROVIDE USER INFORMATION AND A MENU ENVIRONMENT.
<LISTIN>	LIST THE SAS VARIABLES IN ORDER FOR EACH INPUT RECORD.
<DEFINE>	DEFINE THE ATTRIBUTES FOR EACH VARIABLE THROUGH A DATA DICTIONARY LOOKUP AND TO DEFINE THE FORMATS FOR DESIGNATED VARIABLES.
<OUTPUT>	CREATE THE SAS INPUT RECORD FORM AND SAS PROGRAMS TO READ THE INPUT FILE AND CREATE SAS RECORDS AND TO OUTPUT A PERMANENT PROJECT SPECIFIC TEXT LIBRARY.
<ALL>	CALLS <LISTIN>, <DEFINE>, AND <OUTPUT>.
<REVISE>	MAKE CHANGES TO VARIABLE LIST.
<DRUGM>	GENERATE MODULES FOR A DRUG METABOLISM PROJECT.
<END>	EXIT FROM SYSTEM.

SAS/FSP User's Guide, 1982. Edition, SAS Institute, Cary, N.C.

SAS Companion for the VM/CMS Operating System, 1983 Edition, SAS Institute, Cary, N.C.

Contact Author: Jean Bitney  
 A. H. Robins Company  
 1211 Sherwood Avenue  
 Richmond, VA 23220  
 (804)257-2996

**TABLE 2 EXAMPLES OF SYSTEM CALLS**

```
DDCL MET072 LISTIN END
DDCL MET072 REVISE
DDCL DZP018 ALL END
DDCL ISM001 REVISE DEFINE OUTPUT
DDCL ATC015 DRUGM
DDCL LOG MET126 ALL END
```

**REFERENCES**

- Belasco, Barbara and Sues, Suzanne. "SAS: A Data Management Tool Generator," Proceedings of the Eighth Annual SAS Users Group International Conference, SAS Institute, Cary, N.C.
- Hopp, D.I. and Staley, R.W. "Data Management in Clinical Trials Using VM/SP, CMS, and SAS," Proceedings of the Seventh Annual SAS Users Group International Conference, SAS Institute, Cary, N.C.
- SAS Institute, Inc. SAS User's Guide: Basics, 1982 Edition. SAS Institute, Cary, N.C.