

Medical Coding System for Clinical Trials – 21 CFR Part 11 Compliant SAS/AF® Application

Annie Guo, ICON Clinical Research, Redwood City, CA

ABSTRACT

Medical coding in clinical trials is to classify the clinical research data such as adverse events and medications captured during the drug development process. This paper presents a SAS/AF application that addresses the complexity of the task and accommodates standard terminologies such as MedDRA and WHODRUG to ensure the data is interpreted in a consistent manner. It facilitates coding algorithms based on standard dictionaries and proprietary synonyms, auto-encode and manually code/decode verbatim terms, creating predefined reports and user-defined custom reports, and automatically loading dictionary data from ASCII files to SAS. Above all, the application adheres to the requirements of FDA 21 CFR Part 11 and enforces system security as well as maintaining audit trail in a separate database.

INTRODUCTION

The SAS/AF application presented in this paper has been developed in-house on SAS version 8.2 of both UNIX and Windows platforms at ICON Clinical Research. It was originally developed by Mr. Robert Stemplinger for the ICON Mountain View office in California. Soon introduced to other offices, it served our coding studies successfully, but also received users' feedback and there was the need for more diverse functionalities. With the FDA 21 CFR Part 11 put into effect focusing on electronic records, it is crucial to comply with the regulatory requirements for existing computer-based systems. The author began to upgrade the application based on the new system requirements following software development life cycle. The system has completed formal validation.

USER REQUIREMENTS

There are four areas of user requirements.

FDA 21 CFR PART 11 COMPLIANT

Compliance to the 21 CFR Part 11 is required because the coding system modifies clinical data that will be used for regulatory submission by appending the coding information to the data. Some of the regulatory requirements have been in place prior to the upgrade, for example, generating accurate and complete copies of records in both electronic and human readable hard copy format. However, there are two new requirements in order to reach the compliant. First is a secured audit trail that must be maintained by the coding system, and is computer-generated and time-stamped, and that independently records the date and time of operator entries on electronic records. The other is user authentication. In the past the system was accessible to all users who had access permission to the SAS server where the coding system resided. This is now changed so that there is a new layer of user permission at the coding system level; only those on the list of authorized users of the coding system are allowed to log in.

PERFORMANCE

SAS/AF applications run efficiently in general. However, what is unique about this coding system is the functionality of search coding information by running external SAS programs against the dictionary files. Some of the files are fairly large in size, for example, the March, 2008 version of WHODRUG dictionary. So when the SAS programs are retrieving the hierarchy information from various dictionary files, it can noticeably slow down the input-output speed of the coding system. The agreed speed is within 30 seconds for a search algorithm to complete.

CODING

Most of the coding related functionalities have been implemented prior to the upgrade, so the goal is to make the user interface work more efficiently. For example, when a verbatim term is coded, the coding system automatically updates the clinical study data sets with the new coded items, so it leaves no chance of losing work in case of unexpected computer shut-down or not clicking the Save button. In terms of new requirements, the focus is on building a self-contained system to cover all areas of a coding study. For example, in WHODRUG dictionary, with the click of a button, the ingredients of a given drug name are displayed in a pop-up window. Another example, the SAS reporting programs that have been developed in the past are enhanced and organized in the coding system so that everything the users need is in one place. In addition, the reports are available in RTF format as well as Excel format because Excel is a frequently requested report format.

MAINTENANCE

System maintenance is the preparation work to ensure coding assignments are carried out smoothly and

successfully. Examples are setting up system database, importing dictionary data to SAS, and user account maintenance. They are all centralized in the coding system.

DESIGN COMPONENTS

SAS System is FDA's selection as the standard data format, so SAS/AF Software is the chosen application development environment.

SAS/AF FRAMES

The system is developed in SAS/AF software and consists of SAS/AF frames, also referred as screens, as the front-end. The front-end contains SAS/AF objectives to meet a variety of functional objectives. For example, the Table Viewer object is used to display verbatim terms from the clinical study data sets. The Form Viewer object along with Combo Box and List Box objects compose data entry screens that not only display the data from input SAS data sets, but also take in user entries and save back to the SAS data sets.

SCREEN CONTROL LANGUAGE (SCL)

SCL is the core of the coding system that connects the SAS/AF screens. Utilizing various functions, the SCL code controls the data flow among the screens, and activates an appropriate screen based on pre-defined structure for a smooth data flow. For example, when users click the Code button in the main menu, the main menu is immediately deactivated and the DISPLAY function in SCL code activates the coding screen. Other examples are the SET and FETCH functions retrieving data from SAS data sets, the UPDATE function saving changes back to the SAS data sets, and the SUBMIT block interacts with external SAS programs.

SAS DATA SETS

SAS data sets are the ingredients of the coding system; they are digested along the system flow and then presented to users in an understandable manner. They are where the clinical study data and the system data reside. The system data sets include the dictionary data WHODRUG, MedDRA, Costart, ATCVET and VEDDRA. Others examples of system data sets are user account and password information, coding specifications for each study, and programming specifications for each custom report.

EXTERNAL SAS PROGRAMS IN SAS/BASE

Functionalities that are relatively complex in terms of programming are implemented in external SAS programs outside of the frames and SCL. Examples are user login authentication, coding algorithm for each dictionary, updating clinical data sets, system-generated audit trail, reporting, and loading dictionary data to SAS. The purpose of external SAS programs is to minimize the maintenance of the SAS/AF frames and SCL code.

Also implemented in external SAS programs is the search algorithm. The performance of search algorithm is one of the user requirements. To improve the speed of retrieving hierarchy from large dictionary files, SAS indexes are created in primary dictionary data sets. Proc SQL is used in place of Data steps in data processing for efficiency.

PERL SCRIPTS

Perl is the chosen tool for regular expression. It is part of the coding algorithm and search algorithm to find possible synonym names in the dictionary data that match the verbatim terms. The coding system interacts with Perl scripts by X command on UNIX platform, or CALL SYSTEM function on Windows platform.

APPLICATIONS

The practical applications of the coding system cover both the system administrator's and the end users' activities. The system administrator manages thesaurus data, user accounts and other maintenance work. The end users, e.g., the medical coding specialists, carry out the coding assignment thru database lock.

SYSTEM ADMINISTRATION

System administration tasks are performed by users who are specifically granted with administrator rights in the coding system. Their activities include the following.

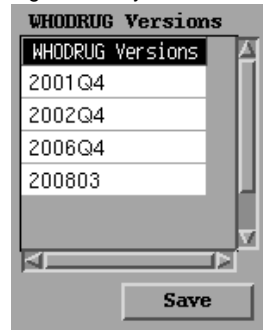
Set Up System Data Sets

First and foremost task is to set up the system maintenance data sets via the Administration screen. The maintenance data sets are used to map to appropriate dropdown lists in other screens for displaying data purposes, or to be read in by external SAS programs as macro variables to perform specific functionalities. Examples are the two data sets that store the dictionaries available in the system (Figure 1.1), and the versions of each dictionary available in the system (Figure 1.2), respectively.

Figure 1.1: System data sets for dictionary types



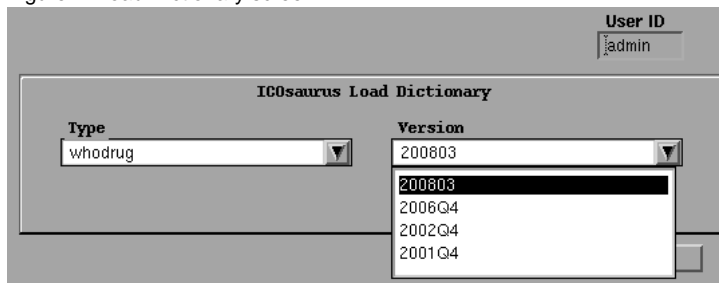
Figure 1.2: System data sets for WHODRUG dictionary versions



Load Dictionary

Next task for the administrator is to save the ASCII data files of the standard dictionaries to the server according to the standard directory organization, and to load them into SAS using the Load Dictionary screen (Figure 2). The screen has two dropdown lists, and their data sources are the two maintenance data sets in Figures 1.1 and 1.2. The administrator clicks on the dropdown lists to select the type and version of the dictionary being loaded. Then click the Run button, and an external SAS program executes to read in the type and the version selected as macro variables, and to import the ASCII files of the selected dictionary version as SAS data sets in a centralized location.

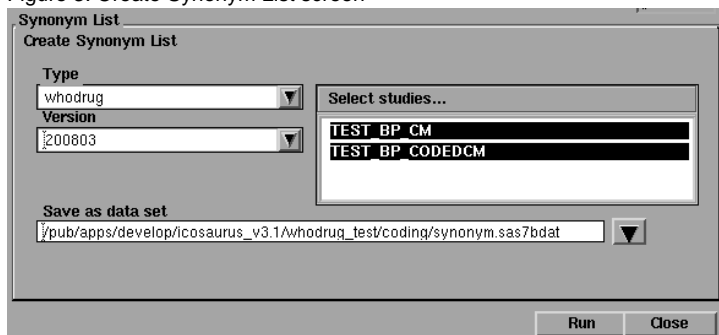
Figure 2: Load Dictionary screen



Create Synonym List

The purpose of synonym list is to achieve consistence in coding across multiple studies. A synonym list can be created from the Create Synonym List screen (Figure 3). First of all the administrator selects the dictionary type and version in the dropdown lists, and then the studies that have been worked on using the same dictionary type and version are displayed in a list box labeled Select Studies. This list box allows for multiple selections, so it is optional whether the synonym list includes data from one or more studies. The coded data sets of the chosen studies in the list box are processed by an external SAS program. The output of the external SAS program is the unique synonym names and their coding data in a SAS data set. The data set is saved in the folder and as the file name specified by the administrator in the text field labeled Save Data Set As.

Figure 3: Create Synonym List screen



Set Up User Account

A medical coding specialist must have an account in order to access the coding system, and this is the administrator's job to set it up. It is done via the Administration-Users screen (Figure 4.1). The screen works like a data entry application in data form layout with a system data set as the backend to store user account information.

To create a new user account, the administrator places the cursor in the combo box labeled User Name, enters a new unique user login name, and types in a new password and re-types to confirm the password. Select the check box Administrator if the new user shall have administrator rights. After all the information is verified, click the Save button, and the new user login information is saved in the user account system data set.

Figure 4.1: Users screen – creating a new user

This screen is also where the administrator maintains the existing user account information. Click the arrow-key on the right of the User Name combo box and a list of the existing user login names is displayed (Figure 4.2). Make a selection, and the selected user's account data are retrieved from the system data set and populated on the screen. Make changes as necessary, or even click the Delete button to remove an old user account. Finally click the Save button to save the changes to the system data set.

Figure 4.2: Users screen – list of existing users

One of the requirements for 21 CFR Part 11 compliant is password encryption. This is implemented thru an external SAS program behind the user account screen that saves and retrieves information from the user account system data set. The way it works is, as the administrator clicks the Save button on the Users screen (Figure 4.1), the SAS program uses a special formula and translates the entered password to a series of symbols that do not have any meaning. So the symbols are actually what are saved on the password variable in the system data set (Figure 4.3). When the administrator clicks the down-arrow and selects an existing user on the Users screen (Figure 4.2), the SAS program reverse-translates the symbols from the data set, and the real password is populated on the screen.

Figure 4.3: System data set for user account information

User Name	User Password	Last Login Date Time	Administrator?
bonnaism	@k**fvm	20MAR2006:17:21:41	Yes
testuser	^n^oymv	11JAN2008:18:21:25	No

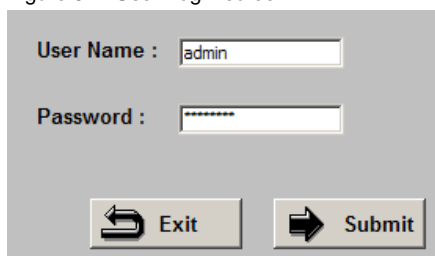
MEDICAL CODING ACTIVITIES

The end users of the coding system are the medical coding specialists. Their activities include the following.

User Login

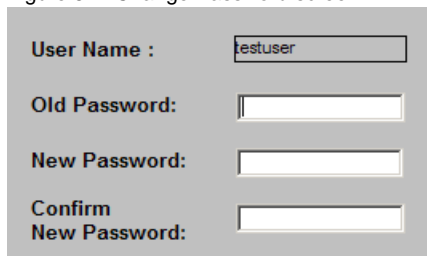
The login screen in Figure 5.1 is a new feature to meet FDA 21 CFR Part 11 compliant. Users enter their username and password and click the Submit button, and then an external SAS program executes to verify the login information. What the SAS program does is to read in the login information, translates the password to symbols based on the special formula for password encryption, and compares with the data in the user account system data set. In addition to password, it also checks whether it has been 90 days since the password was created. If so, the change password screen (Figure 5.2) is displayed and requires users to change password before proceeding.

Figure 5.1: User Login screen



The User Login screen features two text input fields: 'User Name' with the value 'admin' and 'Password' with masked characters. Below the fields are two buttons: 'Exit' with a left-pointing arrow and 'Submit' with a right-pointing arrow.

Figure 5.2: Change Password screen



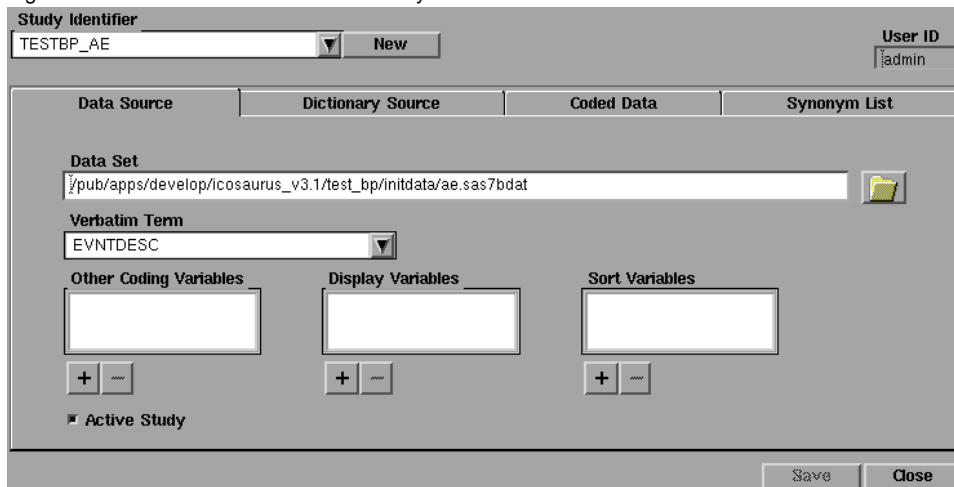
The Change Password screen features four text input fields: 'User Name' with the value 'testuser', 'Old Password', 'New Password', and 'Confirm New Password'.

Defining Coding Study

Defining a study in the coding system is to enter the coding specifications using the Define Study screen. The screen resembles a data entry application with a system data set as the backend. Users begin by clicking the New button to type in a new unique study name in the combo box labeled Study Identifier (Figure 6.1). The coding specifications consist of four parts: data source, thesaurus, coded data, and optional synonym list, each displayed in a tab.

- The Source Data tab (Figure 6.1) contains the folder path and file name of the clinical data set as well as the variable names of the verbatim terms, and for medication coding, indication and route.

Figure 6.1: Data Source tab on Define Study screen



The Data Source tab on the Define Study screen includes a 'Study Identifier' dropdown menu with 'TESTBP_AE' selected and a 'New' button. A 'User ID' field shows 'admin'. The 'Data Set' field contains the path 'y:\pub\apps\develop\icosaurus_v3.1\test_bp\initdata\ae.sas7bdat'. The 'Verbatim Term' dropdown shows 'EVNTDESC'. There are three sections for 'Other Coding Variables', 'Display Variables', and 'Sort Variables', each with a text box and '+' and '-' buttons. An 'Active Study' checkbox is checked. 'Save' and 'Close' buttons are at the bottom right.

- The Dictionary Source tab (Figure 6.2) gathers information about the dictionary to be used for the coding assignment, for example, MedDRA for adverse events and WHODRUG for concomitant medications. In the addition, there is the table viewer labeled Data-Dictionary Variable Mapping where users may edit the standard coding variable names and labels according to sponsor requirements.

Figure 6.2: Dictionary Source tab on Define Study screen

	Original Variable Name	New Variable Name	Variable Label
202	hlgcd	hlgcd	High-Level Group Term Code
203	hlgtnm	hlgtnm	High-Level Group Term Name
204	hlcd	hlcd	High-Level Term Code
205	hltnm	hltnm	High-Level Term Name

- The third is the Coded Data tab (Figure 6.3) to specify the folder path and file name that stores the study's unique and coded verbatim terms.

Figure 6.3: Coded Data tab on Define Study screen

- Finally it is the Synonym List tab (Figure 6.4). This is an optional list of previously coded verbatim terms from other studies to facilitate a new coding study.

After all the information is entered and verified, click the Save button, the entries are saved to the backend system data set.

Figure 6.4: Synonym List tab on Define Study screen

Users can also retrieve the coding specifications of an existing study from the system data set to review or make changes. This is done by clicking the arrow-key next to the combo box labeled Study Identification (Figure 6.1). The restriction of this functionality is if the study has been worked on since its setup, the system does not allow any changes to the specifications. For example, in rare cases if the dictionary version must be changed in the middle of the coding work, that usually means all the verbatim terms are to be re-processed.

Coding Verbatim Terms - Coding Algorithm

The processes of coding verbatim terms are conducted via the Code screen. Users begin by selecting a study in a dropdown list, and immediately an external SAS program executes the coding algorithm. The coding algorithm first match the verbatim terms with the coded items in the coded data set defined for the study and reads in the coding information of those terms. After that if there are terms that remain uncoded, it continues to check against the synonym list, if specified, for identical verbatim terms, and reads in the coding information of the matched terms. Finally if there are still uncoded terms, it attempts to look for identical terms in the dictionary, and retrieve the hierarch information of the matched terms. Especially when using WHODRUG dictionary, only the synonym names with unique ATC codes are included in the dictionary auto-code. The coding results from the three sources are displayed in the table viewer labeled Verbatim Terms (Figure 7.1).

Figure 7.1: Code screen – coding results displayed in Verbatim Terms list box

EVNTDESC	Lower-Level Term Name	Preferred Term Code	Preferred Term Name	High-Level Term Name	High-Level Group Term Name	System Organ Class Name	Coding Level
A SYMPTOMATIC PYURIA							
ABDOMINAL BLOATING	Abdominal bloating	10000060	Abdominal distension	Flatulence, bloating and distension	Gastrointestinal signs and symptoms	Gastrointestinal disorders	A
ABDOMINAL CRAMPING							
ABDOMINAL CRAMPS	Abdominal cramps	10000081	Abdominal pain	Gastrointestinal and abdominal pains (excl oral and throat)	Gastrointestinal signs and symptoms	Gastrointestinal disorders	A
				Gastrointestinal and	Gastrointestinal		

Coding Verbatim Terms – Auto Code vs. Manual Code

The last column in the Verbatim Terms table viewer (Figure 7.1) displays the Coding Level, either “A” meaning auto code, or “M” meaning manual code. Auto-code refers to those terms with exact match in the dictionary and coded as is. Any deviation from auto-code is classified as manual code.

There are several steps to manual code. First users click directly on a verbatim term in the Verbatim Terms table viewer (Figure 7.2). This action triggers a search algorithm for that particular term. The search algorithm is programmed in an external SAS program. The SAS program attempts to find possible matches of the selected verbatim term in the standard dictionary, and populated the results in the list box labeled Possible Matches in Figure 7.2. Alternatively, users may use the text field labeled “Search Text” below the Possible Matches list box to find synonym names only containing specific words to narrow down the number of returns in the list box. Then users click a synonym name in the Possible Matches list box, and another SAS program runs to retrieve the hierarchy information of the selected synonym name from dictionary files, and populates the results in the list box in the lower right area of the screen (labeled “PT, HLT, HLG, SOC” for MedDRA in Figure 7.2). When a correct synonym name is found, users click the Code button on the top of the screen and the coding data along with the coding level “M” is appended to the verbatim term in the Verbatim Terms table viewer (Figure 7.3).

Manual code also covers the re-coding of either previously auto-coded or manually coded terms. The re-coding follows the same steps as above, but a new requirement is to enter a reason for the re-coding in the text field Reason for Re-coding above the Possible Matches list box in Figure 7.3. The text is to be saved in the audit trail for the study for 21 CFR Part 11 compliant purposes.

Figure 7.2: Code screen – processing uncoded verbatim terms

The screenshot shows the 'Code' screen for study 'TEST BP AE'. The 'Verbatim Terms' table is as follows:

EVNTDESC	Lower-Level Term Name	Preferred Term Code	Preferred Term Name	High-Level Term Name	High-Level Group Term Name	System Organ Class Name	Coding Level
A SYMPTOMATIC PYURIA							
ABDOMINAL BLOATING	Abdominal bloating	10000060	Abdominal distension	Flatulence, bloating and distension	Gastrointestinal signs and symptoms	Gastrointestinal disorders	A
ABDOMINAL CRAMPING							
ABDOMINAL CRAMPS	Abdominal cramps	10000081	Abdominal pain	Gastrointestinal and abdominal pains (excl	Gastrointestinal signs and	Gastrointestinal disorders	A

The 'Possible Matches' list includes: Pyuria, Recurrent symptomatic atrial fibrillation, Sterile pyuria, and Symptomatic diffuse esophageal spasm. The search variable is 'Iltm' and the search text is empty. The 'Primary?' checkbox is checked.

Figure 7.3: Code screen – manual code and re-coding

The screenshot shows the 'Code' screen for study 'TEST BP AE'. The 'Verbatim Terms' table is as follows:

EVNTDESC	Lower-Level Term Name	Preferred Term Code	Preferred Term Name	High-Level Term Name	High-Level Group Term Name	System Organ Class Name	Coding Level
A SYMPTOMATIC PYURIA	Pyuria	10037686	Pyuria	Urinary abnormalities	Urinary tract signs and symptoms	Renal and urinary disorders	M
ABDOMINAL BLOATING	Abdominal bloating	10000060	Abdominal distension	Flatulence, bloating and distension	Gastrointestinal signs and symptoms	Gastrointestinal disorders	A
ABDOMINAL CRAMPING							
ABDOMINAL CRAMPS	Abdominal cramps	10000081	Abdominal pain	Gastrointestinal and abdominal pains (excl oral and throat)	Gastrointestinal signs and symptoms	Gastrointestinal disorders	A

The 'Reason for Re-coding' field is empty. The 'Possible Matches' list is the same as in Figure 7.2. The 'Primary?' checkbox is checked.

Coding Verbatim Terms – Save Coding Results

After the initial coding algorithm runs and also every time a term is coded, the coding system runs an external SAS program automatically to save the coding results in the following three data sets.

- First the source data set is updated containing the coding data appended to the corresponding verbatim terms. The updated source data set allows for generating tables/listing of the clinical data for analysis.
- Secondly, the coded data set is refreshed with the unique verbatim terms that have been either auto-coded or manually coded plus the coding information of each term. This data set does not have any subject information. It serves as the source of generating coding reports for review or archive purposes. It is also one of the input data sets when next time the coding algorithm runs. For example, when the clinical source data set is updated with new CRF data, the coded data set associates the previous coding results with same verbatim terms from the new CRF data so that they do not have to be all processed from scratch.
- Finally the audit trail for the study is updated with the changes in the coding of each verbatim term, plus other information according to the 21 CFR Part 11 requirements, such as user id and time stamp when the change

was made, and the reason for re-coding if a term is re-coded.

Reporting

Reporting functionalities are performed using the Reports screen. It is in tab layout and organizes the creation of three types of reports in both RTF and Excel formats: system reports, standard reports, and custom reports.

- The System Reports tab (Figure 8.1) gathers information from one or more studies based on user preferences. The reports aim at providing an overview of the coding studies. For example, the coding status report (Figure 8.2) displays the coding status in terms of the numbers of coded and uncoded terms in Excel format.

Figure 8.1: System Reports tab on Reports screen

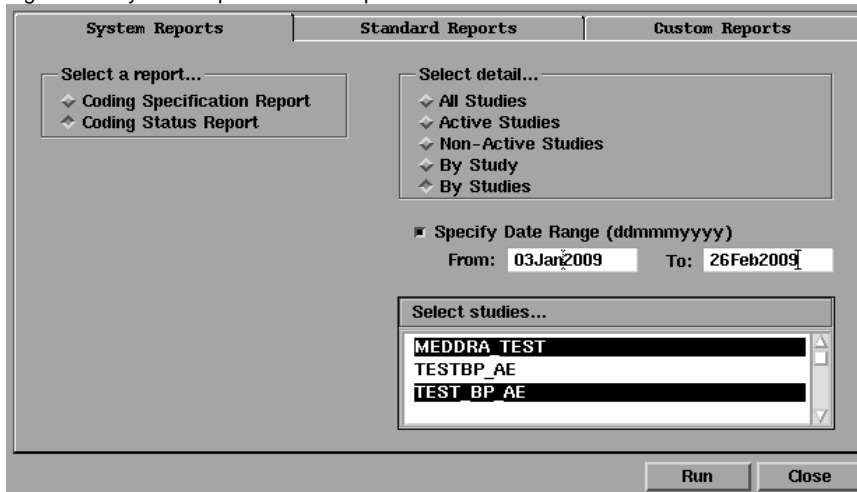


Figure 8.2: System report example – coding status report

ICOsaurus Coding Status Report By Studies
**Date Range : from 03Jan2009 thru 26Feb2009*
Date: 01MAR09

Study ID	Coding Last Modification Date	All Terms				Unique Terms			
		Verbatim Terms	*Coded Terms	Coded Terms (others)	Uncoded Terms	Verbatim Terms	*Coded Terms	Coded Terms (others)	Uncoded Terms
MEDDRA_TEST	24FEB09:17:34:02	308	10	298	0	79	2	77	0
TEST_BP_AE	01MAR09:12:32:44	1900	1428	0	472	600	230	0	370
Total	N/A	2208	1438	298	472	679	232	77	370

System Generated Report 01MAR09 10:41 - By Studies (San Francisco)

- The Standard Reports tab (Figure 9.1) generates study-specific reports according to different coding variables such as verbatim terms and preferred term names. Figure 9.2 is an example in RTF format for the listing of coding by verbatim terms. The general purposes of the standard reports are for reviewing and archiving of the coding studies.

Figure 9.1: Standard Reports tab on Reports screen

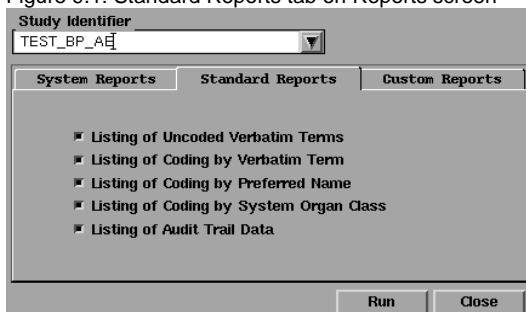


Figure 9.2: Standard report example – listing of coding by verbatim terms

ICOsaurus Standard Report
Listing of Coding By Verbatim Term

Study: TEST_BP_AE
Dictionary: meddra, Version: v11.0

Verbatim Term (CRF)	Lower Level Term	Preferred Term Name	Higher Level Term	Higher Level Group Term	System Organ Class	Coding Level
A SYMPTOMATIC PYURIA	Pyuria	Pyuria	Urinary abnormalities	Urinary tract signs and symptoms	Renal and urinary disorders	Manual Code
ABDOMINAL BLOATING	Abdominal bloating	Abdominal distension	Flatulence, bloating and distension	Gastrointestinal signs and symptoms	Gastrointestinal disorders	Auto Code
ABDOMINAL CRAMPING	Abdominal cramps	Abdominal pain	Gastrointestinal and abdominal pains (excl oral and throat)	Gastrointestinal signs and symptoms	Gastrointestinal disorders	Manual Code
ABDOMINAL CRAMPS	Abdominal cramps	Abdominal pain	Gastrointestinal and abdominal pains (excl oral and throat)	Gastrointestinal signs and symptoms	Gastrointestinal disorders	Auto Code

- The Custom Reports is a new feature for users to create a report according to their specific needs and without actual programming. There are three sub-screens to enter the specifications of a custom report. The first sub-screen is to define the report name, titles and input data set (Figure 10.1). The next is to select the variables from the input data set and define the attributes of the variables to be displayed on the report (Figure 10.2). The third is optional to enter a WHERE clause to subset the data from the input data set. After all the information is entered, users click the Run button (Figure 10.1), and the coding system generates a SAS program based on the specifications, and saves the program in a user-specified folder. And immediately the new SAS program runs automatically to create the custom report. The restriction of this module is that it only allows for one input data set per custom report.

Figure 10.1: Custom Reports tab on Reports screen – define report name, titles and input data set

Figure 10.2: Custom Reports tab – define variables

Figure 10.3: Custom Reports tab – define WHERE clause

AUDIT TRAIL

The functionalities of creating, archiving and reporting audit trail is essential to the 21 CFR Part 11 compliant. The audit trail database is separated from the coding data sets. Figure 11 is the sample audit trail information corresponding to the sample coding report in Figure 9.2. It keeps a chronological log of the coding history. In addition to user name and time stamp and coded items, it maintains the following two variables.

- “Reason in Audit Trail” refers to the action taken in the coding data sets. For example, row #7 in Figure 11 with the reason in audit trail “*Deleted From Dataset*” is saying, because the term was not auto-coded, the coder “admin” manually coded the term and at the same time the blank coding was deleted from the data set. Subsequently row #8 reflects the new coding of the same term in row #7 with Reason in Audit Trail “*Manual Code*”. Another example is rows #11 and #14 where the verbatim terms are auto-coded so the reason in audit trail is “*Auto Code*”.
- Reason Re-Coding is a free text field. When changing an existing coding, users must enter a reason. In Figure 11, the Reason Re-Coding column of rows #9 and #10 is saying the coder replaced the old coding (row #9) with the new coding (row #10), and entered the reason for the re-coding “Coding Correction”.

Figure 11: Audit Trail report

	A	B	C	D	E	F	G	H	I	J	K
1	ICOsaurus Standard Report										
2	Audit Trail Data										
3	Study: TEST_BP_AE										
4	Study Identifier	Dictionary Type	Dictionary Version	Verbatim Term (CRF)	Synonym Name	Preferred Term Name	System Organ Class	Reason in Audit Trail	Reason Re-coding	Date/Time Stamp	Coder ID
5											
6											
7	TEST_BP_AE	meddra	v11.0	A SYMPTOMATIC PYURIA				Deleted From Dataset		14FEB2009:20:44:29 GMT	admin
8	TEST_BP_AE	meddra	v11.0	A SYMPTOMATIC PYURIA	Purpura symptomatica	Purpura	Skin and subcutaneous tissue disorders	Manual Code		14FEB2009:20:44:29 GMT	admin
9	TEST_BP_AE	meddra	v11.0	A SYMPTOMATIC PYURIA	Purpura symptomatica	Purpura	Skin and subcutaneous tissue disorders	Deleted From Dataset	Coding correction	14FEB2009:20:44:55 GMT	admin
10	TEST_BP_AE	meddra	v11.0	A SYMPTOMATIC PYURIA	Pyuria	Pyuria	Renal and urinary disorders	Manual Code	Coding correction	14FEB2009:20:44:55 GMT	admin
11	TEST_BP_AE	meddra	v11.0	ABDOMINAL BLOATING	Abdominal bloating	Abdominal distension	Gastrointestinal disorders	Auto Code		14FEB2009:20:40:36 GMT	admin
12	TEST_BP_AE	meddra	v11.0	ABDOMINAL CRAMPING				Deleted From Dataset		14FEB2009:20:43:50 GMT	admin
13	TEST_BP_AE	meddra	v11.0	ABDOMINAL CRAMPING	Abdominal cramps	Abdominal pain	Gastrointestinal disorders	Manual Code		14FEB2009:20:43:50 GMT	admin
14	TEST_BP_AE	meddra	v11.0	ABDOMINAL CRAMPS	Abdominal cramps	Abdominal pain	Gastrointestinal disorders	Auto Code		14FEB2009:20:40:36 GMT	admin

CODING STUDY LOCK

After completion of a coding study, users uncheck the check box labeled Active Study on the Define Study screen (Figure 6.1) and click the Save to save the information to the coding specification system data set. This action removes the study name from the list of studies available for coding on the Code screen (Figure 7.2), and as a result no further coding work is allowed and the coding study is locked in the coding system.

PERFORMANCE

Some of the dictionary data contains big data sets, and it affects the speed of the coding system when retrieving the hierarchy information of the synonym names from the data sets. The coding system uses the following methods to achieve better performance.

- Proc SQL is used to merge multiple data sets efficiently.
- Compact SAS data sets are created containing a subset of the information from big primary data sets. For example, for WHODRUG, there is a data set including only the drug names with unique ATC code. This facilitates the auto-code process when using WHODRUG dictionary.
- SAS Indexes are created in primary dictionary data sets, for example, the compact data set mentioned above for WHODRUG.

CONCLUSION

The coding system is formally validated and complies with the latest regulatory requirements as well as the new features based on users’ feedback. It has been developed following software development lift cycle with complete documentation in archive.

DISCUSSION

The system can accommodate additional dictionaries. To achieve that, a few frames such as the Code screen and the SCL code need to be updated, and also external SAS programs for the coding algorithm of the new dictionary.

ACKNOWLEDGMENTS

The author would like to extend a special thanks to Robert Stemplinger for his review of this paper, as well as his ongoing support for progress, creativity and innovation.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Annie Guo

ICON Clinical Research

Redwood City, CA, USA

E-mail: annie.guo@iconplc.com, annieguo@yahoo.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. © indicates USA registration.

Other brand and product names are trademarks of their respective companies.