DIRECTIONS IN REGISTRY MANAGEMENT SOFTWARE DEVELOPMENT: The DHCP Oncology Package

Michael L. Henderson
Department of Veterans Affairs
Washington Information Systems Center
8403 Colesville Road, Suite 200
Silver Spring, MD 20910
(301) 427-3700
henderson.michael@forum.va.gov

Abstract

The Department of Veterans Affairs (VA) Oncology/Tumor Registry Package is an evolutionary registry management tool within the VA Decentralized Hospital Computer Program (DHCP). The package, written in ANSI standard M, combines a VA Fileman data base with a rule base and statistical analyzer. It automates the chore of gathering cancer data for existing hospital patients and contains up-to-date implementations of cancer staging algorithms. This functionality allows the package to support medical center cancer programs that meet the standards established by the American College of Surgeons (ACOS).

As is the case throughout DHCP, this package design avails itself of the significant development benefits that accrue in an M/Fileman/Kernel environment. The accelerating pace with which specifications are evolving make it essential that programming constructs be used that facilitate timely creation of package versions and modifications. We shall discuss the development of Oncology Version 2.1 and its integration into the new and existing DHCP tool sets. Additionally, we shall steal a glimpse into the future of Oncology development and how the package is preparing for the object paradigm.

Introduction

Since its introduction in 1988, the DHCP Oncology/Tumor Registry Package has sought to incorporate the database features necessary to support tumor registries at all 172 VA Medical Centers. Its initial purpose was to serve as a registrar's data recording tool within DHCP. This allowed permanent computerized storage and retrieval of tumor data within the guidelines set by the American College of Surgeons [3,4,5,9] but did not significantly reduce the clerical task load of the registrar.

Although referred to as Oncology, the package as it has been developed to date is specifically a tumor registry package. The tumor registrar's goal is the accurate collection and characterization of tumor data, as compared with the goals of the clinician charged with the direct treatment of the cancer patient. Clinical care of the oncology patient is dealt with in other modules of DHCP, including portions of the Medicine, Surgery, Laboratory, Radiology, and Patient Information Management Service (PIMS, formerly Medical Administration Service) packages. As we shall be discussing, it is essential that these packages interact with one another to support the total care environment.

Subsequent releases of the package have taken greater advantage of the data integration available within DHCP, the relational file constructs and user help features of VA Fileman, and the enhanced structure of M for performing computational tasks. Later versions of Oncology have included an automatic casefinding mechanism that gathers data from multiple DHCP modules to track suspected cancerous lesions [1]; a set of algorithms to derive accurate cancer stage from data entered by the registrar [2]; and invocations available from within VA Fileman to perform cross-referencing and multiple-file maintenance to the degree necessary to ensure maximal data integrity [6]. In addition, routines have been developed that process subsets of data produced by Fileman to produce tumor registry statistical reports [7].

As Version 2.0 was released to the field, VA clinical developers continued to seek new ways to incorporate M and Fileman features into a viable, enhanceable package. The necessity for a proactive development approach is dictated by the dynamic environments in which VA tumor registries find themselves operating:

- Data gathering and analysis needs are ever changing. Since the publication of the first edition of the *Manual on Staging of Cancer* in 1977, three full editions have followed, and a fifth (and possibly a sixth) edition is anticipated in the next few years. Each succeeding edition contains significant refinements and revisions of the characterization and staging mechanisms. Additionally, a new revision of the ACOS *Data Acquisition Manual* is due to be released in 1995. New revisions are also forthcoming for the *SEER Extent of Disease Codes and Coding Instructions*. [3,4,9]

- The DHCP software environment is being transformed. A subgroup of ANSI X11 is developing a MUMPS Windowing Application Programmer Interface (MWAPI), which completely alters the way in which data is displayed and entered.

- M and Fileman systems developed by VA are expected to evolve in such a way as eventually to interact seamlessly with specialized commercial systems [8].

The Market for Tumor Registry Software

The collection of cancer data is performed by a number of software packages. In addition to DHCP Oncology/Tumor Registry, two off-the-shelf packages—CANSUR/NET, produced by ACOS, and MRS (Medical Records System), a commercial application—have been in use for some years in tumor registries in many DVA medical centers. Medical centers that are considering conversion from these latter packages have come to depend on the functionality they provide. While all of these packages were designed to perform registry functions, each has some relative advantages.

CANSUR/NET is a package designed by the American College of Surgeons. One of its main functions is to perform the collection of data specifically required in the College's annual Call for Data. It is a standalone PC product. The state of California uses an adaptation called C'NET in its central registry.

MRS (Medical Records System) is a commercial package that also runs standalone on PCs. It has had considerable popularity, particularly in the Northeast, because of its reporting interface and high level of user support. However, these features also come with a high price, and the necessity to reallocate VA resources in an era of budget cutting made it difficult to justify the continued expense of this product.

DHCP Oncology is the newest of the three, which is its chief disadvantage although a temporary one. The advantages it offers are considerable; two merit particular mention. The package provides automatic, customizable casefinding through integration with the DHCP PIMS, Radiology, and Laboratory packages. Additionally, data provided by the system are available to all authorized DHCP users. By incorporating such functionality, the package has become as much a registry management tool as a data collection tool, and this trend is anticipated to continue.

In addition, automatic staging is supported for cases diagnosed after 1988. ACOS reporting is facilitated through a PC interface with whatever platform the package is running on. A statistical module is provided to assist the registrar in reporting of cancer survival and outcomes.

Aspects of Cancer Staging

The staging of cancer is a complex problem that can be further complicated by the location of the cancer and the amount of data available. In most cases it is necessary to know the size of the tumor, its *topography* or precise anatomical location, the *histology* or cell architecture, the degree of regional lymph node involvement, and whether or not distant metastasis is present. For certain cancers, it is also necessary to know the patient's age and the grade of cell differentiation (the range of which is dependent upon *histo-morphology*, the sub-classification of the cell itself).

Once these factors are known, it is possible to determine the staging group into which the cancer falls. As of the latest revision of the AJCC *Manual on Staging of Cancer*, there were more than fifty staging groups, some of which contain multiple algorithms. Staging can further differ based upon whether the basis of classification is *clinical* (based on evidence acquired before treatment, such as X-ray examination) or *pathologic* (based on clinical evidence modified by tissue evidence from surgery, biopsy, or autopsy). Certain anatomic sites, such as those within the head and neck, require use of a classification hierarchy in order to arrive at the correct staging. [3]

As is evident, the problem of classification of cancerous lesions resists superficial tabular solutions. This comes as no surprise to the clinical practitioner but nevertheless complicates the task of the clinical software developer in generating the correct data collection and characterization methodology.

The M/Fileman Integrated Solution

Version 2.1 of DHCP Oncology/Tumor Registry includes refinements to all of the data characterization routines. The package implements a decision hierarchy in determining appropriate domains for such data as regional and distant lymph node extension.

All tables necessary for data characterization reside in Fileman files. However, depending upon the factors cited above, the particular table to be chosen may be associated with topography, histology, or staging group. Many topographies, histologies, or staging groups may be associated with a single table, and logic for the associated file takes this into account. As Figure 1 illustrates, complicated M algorithms are required to map the lesion data to the appropriate tables. The algorithms are referenced in the input transforms of the data dictionary for the Oncology Primary file.

Once the data are gathered and characterized, tests are applied for consistency. Upon validation that all necessary items are in place, appropriate staging algorithms are invoked that apply logic from the third or fourth edition of the *Manual on Staging* (Figure 2). An effort is made to include all necessary processing code in each staging module, rather than burying code in the calling modules, so that it will be a relatively simple matter to add new logic for those algorithms that change in future editions. The goal here is to reduce the number of 'surprises' for developers and testers of future editions of the package.

An Eye on the Future

With the migration of M and Fileman into Windows and other graphical environments, it will be necessary to code applications to a variety of input/display devices. This objective can be partially met today by building small libraries of modular utility code (see Figure 3 for an example) that in turn contain calls to Fileman utilities such as the Reader and (in future) Writer, which can handle the cosmetic portion of the interface without requiring extensive modification of the application code. As Fileman grows to support more sophisticated screen displays and graphical user interfaces, a relatively minor alteration of the utility function should be all that is needed. The program modules that invoke the utility can remain unchanged.

Conclusion

As the reader may by now have gathered, the constructs used here do not attempt to impose a radically new style of programming upon the package. The principles by which development has proceeded are grounded in the philosophy of structured, modular programming that have matured over the last two decades. An attempt is being made to move toward principles of object orientation, particularly with respect to the encapsulation of data and function. These practices should serve to make the Oncology Package not only a useful tool in itself, but an example to be emulated as other registry management tools (*e.g.*, for diabetes) are developed within DHCP.

As M evolves toward the object paradigm and its developers move more comfortably within the object world, complex packages such as Oncology can incorporate additional functionality as needed with little risk that relatively sudden changes or additions might compromise other portions of the package. Optimal adaptation to this evolution will insure the investment that the nation has made in VA's clinical information resources. Our perpetual goal is to enhance the utility of those resources to the customers that VA and DHCP serve.

Acknowledgement

Special thanks is due to Susan Richie of VA's Washington Information Systems Center, whose hard work and dedication have laid the foundation for the development of the automated tumor registry within VA.

References

1. Richie, Susan, "Automatic Casefinding of Cancer Integration of Databases in the VA DHCP Automated Tumor Registry," *M Computing* 1:3 (June 1993), 146-152.

2. Richie, Susan R., "Automatic TNM Staging of Cancer Using the VA-DHCP Oncology-V2 Tumor Registry," *M Computing* 1:3 (June 1993), 153-162.

3. American Joint Committee on Cancer, *Manual for Staging of Cancer*. Philadelphia: J. B. Lippincott Company, 1992.

4. American College of Surgeons Commission on Cancer, *Data Acquisition Manual*. Chicago: American College of Surgeons, 1988.

5. American College of Surgeons Commission on Cancer, *Cancer Program Manual*. Chicago: American College of Surgeons, 1991.

6. Richie, Susan Roberts, and Johnson, Martin E., "File Manager Portals," *MUG Quarterly* 21:3 (June 1991), 94-96.

7. Richie, Susan Roberts, and Marciniak, Thomas A., M.D., "File Manager and Statistics," *MUG Quarterly* 21:3 (June 1991), 97-100.

8. "Long Range Strategy for Inter-Application Communication," VA monograph, September 1993.

9. American Joint Committee on Cancer, *SEER Extent of Disease Codes and Coding Instructions*. Chicago: American College of Surgeons, 1988.

```
GETLIST(ONCOIX) ;   find the extension/lymph node list that applies to primary ONCOIX
      ;   could be in 164, 164.1, or 164.2
      N OG,OP ;   global and record number
      N ONCOER ;   error code
      N OS,O2 S OS=$P(^ONCO(165.5,D0,0),U),O2=$G(^(2)) I O2="" S ONCOER="ID" ;   cancer ID must exist
      I $G(ONCOER)="" N ONCOT S ONCOT=$P(O2,U) I ONCOT="" S ONCOER="TOP" ;   topography must be
valid
      I $G(ONCOER)="" N ONCOM S ONCOM=$P(O2,U,3) I ONCOM="" S ONCOER="HIST" ;   histology must be
valid
      IF $G(ONCOER)="" D ;   we have the necessary information, get the list
      . IF $$MELANOMA^ONCOU55(D0),$P($G(^ONCO(164,ONCOT,0)),U,15) S OG=164.2,OP=39 ;   malignant
melanoma of the skin?  yes.
      . ELSE  D ;   no. check further
      .. N ONCOMP S ONCOMP=$P($G(^ONCO(164.1,ONCOM,0)),U,3) ;   morphology extension/lymph
node encoding pointer
      .. I ONCOMP S OP=ONCOMP,OG=164.1 ;   characterized by morphology
      .. E  I ONCOT=67770&(OS=62!(OS=63)) S OG="164.2",OP=OS ;   lymph nodes
      .. E  I (OS=35)!(OS=39)!(OS=40) S OG=164.2,OP=OS ;   formerly also for OS=38
      .. E  S OP=$P($G(^ONCO(164,ONCOT,0)),U,3),OG=$S(OP="":164.2,1:164),OP=$S(OP="":OS,1:OP)
      .. Q
      . ;END IF
      . ;
      . Q
      ELSE  D ;   there is an error
      . W !
      . I ONCOER="ID" W "Cancer ID is required!"
      . E  I ONCOER="TOP" W !,"Topography code is required!"
      . E  I ONCOER="HIST" W "Histology code is required!"
      . W *7,!
      . Q
      ;END IF
      ;
      N ONCOOUT S ONCOOUT=$G(OG)_U_$G(OP) S:$G(ONCOER)'="" $P(ONCOOUT,U,3)=ONCOER
      Q ONCOOUT
```

FIGURE 1: This abstracting function examines topography (ONCOT), histology (ONCOM), and site group (OG) and returns the Fileman file number of the appropriate table for lookup of SEER extension and lymph node codes.

```
30    ;
      K SG
      D @$S(ONCOYR:"TST",1:"TS3")
      I '$D(SG) S SG=9 ;    unstageable
      Q
      ;
TST   ;TESTIS
      IF M=1 S SG=3 ;    distant mets?  yes
      ELSE  IF M=0 D  ;    no. check tumor & local nodes
      . I N=0 S SG=$S(T="IS":0,1:1) ;    local nodes?  no
      . E  S SG=2 ;    yes
      . Q
      ;END IF
      ;
      Q
      ;
TS3   ;3rd Edition
      IF M=1 S SG=4 ;    distant mets?  yes
      ELSE  IF M=0 D  ;    no. check tumor & local nodes
      . IF (N=2)!(N=3) S SG=4 ;    local nodes?  yes...
      . ELSE  IF N=1 S SG=3 ;    ...yes...
      . ELSE  IF N=0 D TS3TUM ;    ...no. check tumor
      . ;END IF
      . ;
      . Q
      ;END IF
      ;
      Q
      ;
TS3TUM  ;    TESTIS - 3rd edition - check tumor if no local/distant mets
      I T="IS" S SG=0
      E  I (T=1)!(T=2) S SG=1
      E  I (T=3)!(T=4) S SG=2
      Q
```

FIGURE 2: As shown in this routine excerpt, staging functions incorporate extensive block structuring. Differentiation logic seeks to efficiently process cases while keeping the code maintainable.