Sharing MLM's: An Experiment between Columbia-Presbyterian and LDS Hospital

T. Allan Pryor, Ph.D.
Department of Medical Informatics, LDS Hospital/University of Utah
Salt lake City, UT 84143

George Hripcsak, M.D.
Center for Medical Informatics, Columbia-Presbyterian Medical Center
New York, NY 10032

The use of Arden syntax for development of shareable medical logic modules (MLM's) has developed as an ASTM standard. To test the feasibility of sharing MLM's between institutions a study was conducted between Columbia-Presbyterian Medical Center and LDS Hospital. In this study seven MLM's clinically executing at Columbia-Presbyterian were used to test the sharing potential of the Arden syntax. The study was limited to measuring the modifications necessary to make executable at LDS Hospital the shared MLM's. Because of the site specific nature of the data variables, multiple modifications were required. Three classes of modifications were necessary. The simplest involved only data variable mappings. The other classes required either minor modifications to the logic or relatively major modifications. Over 50% of the modifications were in the minor or major classes. While the sharing of decision logic was possible and facilitated by the use of the MLM's at the two sites, the absence of standard medical vocabularies limited the utility of the MLM as a mechanism for directly sharing medical knowledge.

INTRODUCTION

Creation of clinical knowledge bases is an active research interest of almost every center of Medical Informatics. Unfortunately this usually results in knowledge bases that are usable at only the site where the knowledge base was created or through duplication of the system on which the knowledge base was created. The sharing of well documented and validated knowledge bases could markedly increase effectiveness of research done at these isolated centers by making available to all the success of individual researchers.

One effort undertaken to assist in sharing medical knowledge has been the creation of a standard language for description of medical knowledge. This effort resulted in the definition of the Arden syntax which has been published by ASTM as a voluntary standard. The Arden syntax is based on the early work of HELP at LDS Hospital and CARE at Regenstrief Hospital and describes a modular representation of medical knowledge designed to support small independent sets of medical knowledge in the form of discrete logic statements. These modules known as Medical Logic Modules (MLM's) have the advantage of being potentially useful even if only a single module exists on a system. Because of the MLM's independence, the set of modules can easily be added to or modified without affecting the performance of other MLM's running on the system.

To test the ease of sharing MLM's between institutions an experiment was carried out between Columbia-Presbyterian Medical Center and LDS Hospital. In this experiment 7 MLM's clinically running at Columbia-Presbyterian Medical Center were transferred to LDS Hospital. The MLM's were then modified to execute clinically on the HELP system at LDS Hospital.

Before an MLM can be implemented at a second site, two issues need to be resolved at the receiving site. The first is simply to have the appropriate medical experts at the receiving site agree to the discrete logic statements as contained in the MLM. The second issue is a technical one. In writing an MLM, the author must define the logic in terms of data elements that are supported on his/her information system. Since there are no standards for description of data elements, similar data elements may have a much different definition or structure at the

0195-4210/92/$5.00 © 1994 AMIA, Inc.

receiving site. Therefore, to implement the received MLM, the receiving institution is required to modify the MLM to utilize corresponding data elements from their local medical vocabulary. To modify the MLM, the receiving institution must verify that the data elements exist in their database. If corresponding data elements do exist, the receiving site then makes the appropriate mappings between data elements as defined at the sending site and those available at the receiving site. After making the appropriate changes in data elements the receiving institution must check the logic of the MLM to determine if any logic changes are necessary to accommodate the site specific changes to the data elements. This paper outlines the modifications to the MLM's that were required for LDS Hospital to implement MLM's received from Columbia-Presbyterian Medical Center. No attempt was made to measure the acceptability of the logic by the medical experts at LDS Hospital.

SYSTEM ENVIRONMENTS AT THE TWO SITES

The Hospital Information System's (HIS's) at Columbia-Presbyterian and LDS Hospital are quite distinct. The IBM Patient Care System is the basic HIS used at Columbia -Presbyterian. The clinical database uses the IBM relational DB2 for storage of their clinical database. The laboratory data is received by the IBM from an in house developed Laboratory Information System. The laboratory data is encoded with a set of codes representing the test performed and a set of values that are either ASCII text or numeric. MLM's written at Columbia-Presbyterian are compiled to a set of pseudo codes that are interpreted on the IBM mainframe. The MLM's shared in this experiment were written to be data driven, i.e., executed on receipt of data into the patient's clinical database.

HELP is the clinical HIS at LDS Hospital and uses a multiprocessor Tandem computer as the main platform for execution of all clinical applications. The HELP database referred to as PTXT is HELP specific. The Laboratory system at LDS Hospital is a commercial LIS developed by the Lab Force Company. All data stored in HELP from the laboratory system are PTXT coded with numeric or ASCII values. MLM's written at LDS Hospital are compiled to the HELP application language PAL. A PAL module has a one to one correspondence to its associated MLM. As with the Columbia-Presbyterian system, execution of the MLM's is implemented in the data driven mode.

DESCRIPTION OF MLM'S

Seven MLM's were used in the study. Each of these MLM's is in clinical operation at Columbia-Presbyterian. Table 1 lists the titles of the seven MLM's. From the titles that are listed, it would appear that two of the MLM's are redundant. In particular there appears to be two MLM's that are used to screen for hypokalemia and two to screen for Hepatitis B surface antigen in mothers. In fact these are slightly different in their logic and their existence is a direct result of the philosophy behind the creation of an MLM. The issue is whether to create one complicated MLM which in the case of the hypokalemia screen would be triggered by both the clinical storage of a digoxin level and the storage of a potassium level, or as the authors at Columbia-Presbyterian have chosen, create two simpler MLM's, one being triggered by storage of the digoxin level and the other by the potassium level. The effect of site specific philosophy in the creation of MLM logic is another area whose effect on the ability to share MLM's needs to be evaluated but was not considered in this study. No attempt was made to modify the MLM's to support a different philosophy of MLM creation, but only those modifications required to make the MLM operational at LDS Hospital were made.

Table 1
Titles of MLM's from Columbia Presbyterian Medical Center

MLM1 - Screen for hypokalemia with digoxin therapy (triggered by potassium storage)
MLM2 - Screen for hypokalemia with digoxin therapy (triggered by digoxin level storage)
MLM3 - Fractional excretion of sodium
MLM4 - Screen for positive hepatitis B surface antigen in mothers
MLM5 - Screen for positive hepatitis B surface antigen in newborn's mothers
MLM6 - Screen tuberculosis cultures for positive or invalid results
MLM7 - Screen for worsening renal insufficiency based upon Serum Creatinine

All of the MLM's in this study were simple alerting MLM's that are data driven by the storage of clinical data. The result of each of the MLM's was an alert that was sent to the alerting system of each system. In the case of Columbia-Presbyterian this was both a message in the patient's computer record and a message in their

research log. At LDS Hospital only the message in the patient's record was created since no similar research log existed.

DESCRIPTION OF THE MODIFICATIONS MADE TO THE MLM'S

The simplest form of modification made is illustrated in the following example.

```
Columbia-Presbyterian Syntax
admission := event {0,09000101;0,09000106};

LDS Syntax
admission := event{admit which is
        49.1.25.2.8.3};
```

Example 1
Syntax Differences Between LDS and Columbia-Presbyterian for defining a simple MLM variable.

In this example the data elements within the curly brackets are changed to the corresponding codes at LDS Hospital. The data elements at both sites consist of codes and only the syntax of associating those codes with the knowledge variables is changed. Note in this example that in some instances multiple codes may be used at one site that are mapped to a single code at the other site. With modifications of this type no logic changes are required and the changes are easily accomplished by understanding the appropriate mapping of the database elements between the two sites.

Example 2 illustrates a more complex understanding required between the database models at two sites.

```
Columbia-Presbyterian Syntax
service := read last {'evoking','dam'=
        "GYDAPMP",'constraints'="I";
        "HCASEX"; "HADMSVCL"};

Logic using service
If service is null or service <> "obs" then
        conclude false;

LDS Syntax
service := read last {class WHICH IS
        1.0.1.0.16.76.8.8};

Logic using service
If service <> 2 then  conclude false;
```

Example 2
Difference in Syntax between LDS and Columbia-Presbyterian resulting from Medical Vocabulary Structure

In this example we see that because of the data models at Columbia-Presbyterian logic has been included in the read statement to ensure that the appropriate value will be assigned to the variable service, whereas at LDS the mapping is to a simple coded data element. The differences in the data models resulted in the minor changes seen in example 2. In this example, the modifications were due to the ASCII versus numeric structure of the variable in the different database representations. Common among the changes required to implement the MLM's at LDS Hospital were minor changes in logic statements to reflect the different data representations of the two systems. Usually, the changes reflected the use of ASCII versus coded data elements.

The third example illustrates a case where more extensive logic modifications were required. In this example, differences in clinical laboratory instrumentation between the two sites necessitated the changes. At Columbia-Presbyterian different instruments are used for different chemistry panels. The value a variable is slightly different depending on which chemistry panel was performed. Because of the differences in values, logic was included in the Columbia-Presbyterian MLM to correct the creatinine value based on the panel which was measured. AT LDS Hospital, the data variables are measured from the same instrument regardless of the test panel, thus eliminating the need for logic to correct the variables based on instrument. This example illustrates a dependence of an MLM not only on a site specific data model, but also a dependence on other site specific factors that need to be understood before conversion of the MLM can be completed.

```
Columbia-Presbyterian Syntax
if current_panel = "CC000002" then/*panel is a
        SMAC*/
  corrected_creatinine := current_creatinine-0.15;
else                            /*panel is a chem7*/
  corrected_creatinine := current_creatinine;

LDS Syntax
no syntax required since values come from same
        instrument regardless of panel
```

Example 3
Differences in LDS and Columbia-Presbyterian Syntax resulting from differences in database structures

The above examples illustrate the classes of changes that were necessary to implement the

MLM's from Columbia-Presbyterian on the LDS HELP system. Table 2 lists the number changes made in each of three categories to translate the 7 rather simple MLM's.

Table 2
Frequency of modifications to MLM's

Type 1 modifications (data mappings only)	22
Type 2 modifications (minor logic modifications)	18
Type 3 modifications (major logic modifications)	3

Table 3 gives the number of changes made to each MLM in each of the Arden syntactical slots. As seen in the table every data statement naturally had to be changed. Likewise all of the action statements changed since alerts were recorded differently on each of the systems. The most interesting statistics were the number of changes made to the logic statements to accommodate the differences in databases. As reported in Table 3 the number of changes to logic statements ranged from 0 in the Fractional excretion of sodium MLM to 8 in the MLM for screen renal insufficiency.

Table 3
Number of statements/Number of changes to MLM defined slots

	Slot label		
	Action	Data	Logic
MLM1	5/5	6/2	2/2
MLM2	5/5	5/1	2/2
MLM3	6/6	3/0	2/2
MLM4	5/5	5/3	2/2
MLM5	7/7	6/3	2/2
MLM6	4/4	4/3	2/2
MLM7	10/10	20/8	2/2

DISCUSSION

While the development of a common syntax for sharing medical logic is a first step, the results of this experiment emphasize the need for a common medical vocabulary. 50% of the changes made to the MLM's were changes not only to the data variable definitions, but involved some modifications to the logic of the MLM. In the logic slot alone 20 of the 49 logic statements in the 7 MLM's needed to be modified. The magnitude of the problem is unclear since the changes that were made were mechanical and did not involve additional knowledge engineering. Had a common vocabulary existed, however, only the one MLM involving changes due to the differences in the measurement of variables from the instruments used would have needed logic modification.

The conversion process itself, while fairly mechanical, required several tasks. The first was to understand the structure of the clinical database where the MLM was created. This involved several sessions with those at Columbia-Presbyterian inquiring about the meaning of the syntax within the curly brackets. The second task was to define the appropriate map between the Columbia-Presbyterian data variables and those at LDS. In the simplest case this was a one to one mapping, but in a majority of the cases it was a many to one mapping and more importantly an understanding of the structural differences in the data variables of the two systems. The final task was to scan the logic to understand how the logic was dependent on the structure of the variables. Modifications to the logic were limited to conserving the same concept and using new logic only to reflect the different LDS structure. Since no test data set was available from Columbia-Presbyterian to compare the performance of the modified MLM's at LDS, it is unclear that the two sets of MLM's would have given precisely the same results. The problem, however, in using a test database for validation of the MLM's is problematic, since the same problems occur with the data set as with the MLM's. That is, to use the test data set on the new MLM's, it would be necessary to convert the Columbia-Presbyterian data structure to the LDS data structure. This, however, could sufficiently modify the data set that it no longer was reflective of the data set used at Columbia-Presbyterian.

In summary, the sharing of medical knowledge remains a difficult problem. The use of Arden Syntax while helpful, has not overcome some of the time consuming requirements to share clinically useful decision logic. While this effort should continue, its success may be delayed until a common medical vocabulary together with conversion routines that convert site specific databases into this standard vocabulary exist. Our experience suggests that standards can be of great assistance in sharing the work of many, but the routine sharing of medical knowledge may be delayed until common standards exist not only in the description of the logic but in all aspects of the medical information system.

References

[1]. Pryor TA, Gardner RM, Clayton PD, Warner HR. The HELP system, In Blum BI ed., Information Systems For Patient Care, New York: Springer Verlag, 1984;109-128

[2]. Kuperman GJ, Gardner RM, Pryor TA. HELP: A Dynamic Hospital Information System. Springer-Verlag Inc., New York, NY 1991.

[3]. McDonald CJ. Action-Oriented Decisions in Ambulatory Medicine. Chicago: Year Book Medical Publishers, 1981.

[4]. McDonald CJ. Protocol-Based Computer Reminders, the Quality of Care and the Non-Perfectibility of Man. N Eng J Med 1976; 295:1351-1355.

[5]. McDonald CJ, et al. The Regenstrief Medical Record System: 20 Years of Experience in Hospitals, Clinics, and Neighborhood Health Centers. MD Comp 1992; 9(4):206-217.

[6]. Clayton PD, Pryor TA, Wigertz OB, Hripcsak GM. Issues and structures for sharing knowledge among decision-making systems: The 1989 Arden Homestead Retreat. Kingsland LC ed., Proc 13th SCAMC, IEEE Comp Soc Press, 1989;116-121.

[7]. Clayton PD, Pryor TA, Wigertz OB, Johnson SB, Hripcsak GM. Sharing medical knowledge for automated decision-making. Greenes RA ed., Proc 12th SCAMC, IEEE Comp Soc Press, 1988;591-594.

[8]. Wigertz OB, Clayton PD, Hripcsak G, Linnarsson R. Knowledge representation and data model to support medical knowledge base transportability. In Talmon J, Spiegelhalter J, System Engineering in Medicine, Heidelberg: Springer Verlag, 1989.

[9]. Yasnoff WA. Technology transfers of medical decision assistance systems. Hammond WE ed., AAMSI Congress 89, AAMSI Publishers, 1989;103-107.

[10]. HELP Frame Manual, Version 1.6. Salt Lake City: LDS Hospital, 1989.

[11]. McDonald CJ, Blevins L, Tierney WM, Martin DK. The Regenstrief Medical Records. MD Computing 1988;5(5):34-47.

[12]. Standard Specification for Defining and Sharing Modular Health Knowledge Bases (Arden Syntax for Medical Logic Systems), Destination E 1460 - 92, ASTM, 1916 Race Street, Philadelphia, PA 19103-1187.

[13]. Anderson RK, Clayton PD, et al. Integrated Academic Information System (IAIMS) Implementation at Columbia-Presbyterian Medical Center. New York: Columbia-Presbyterian Medical Center, 1988

[14]. Shwe M, Sujansky W, Middleton B, Reuse of Knowledge represented in the Arden Syntax. Frisse ME ed., Proc 16th SCAMC, IEEE Comp Soc Press, 1992;47-51.