Translating National Childhood Immunization Guidelines to a Computer-Based Reminder Recall System within an Immunization Registry

Dongwen Wang, Robert A. Jenders, MD, MS, Balendu Dasgupta, MS
Department of Medical Informatics, Columbia University
New York, New York 10032

To translate national childhood immunization guidelines to a computer-based reminder recall system, hierarchical system architecture design and combined approach of tabular and procedural knowledge representation are taken. Nested branches with hierarchical combinations of single antecedent variables are used to avoid logical incompleteness, redundancy and inconsistency. Mapping to the local electronic medical vocabulary is implemented to facilitate the integration with the local information system architecture. 26 second-level modules with 195 original branches and 121 final branches after pruning are encoded. 99.67% of the reminders are confirmed to be correct by SQL query.

INTRODUCTION

Immunization, Guidelines and Information System

Although the immunization coverage rate in the United States is increasing as a whole[1,2], children in inner city and low-income communities are still underimmunized[3,4].

Clinical guidelines are used in preventive care to improve the compliance with health care standards[5]. Immunization guidelines are used to improve the immunization coverage rate[6]. To address this need, the Centers for Disease Control and Prevention (CDC) have been publishing the General Recommendations on Immunization and its modifications every few months since 1994[7].

Studies show that clinical decision support systems (CDSS) based on clinical guidelines can improve guideline compliance[8,9]. Development of immunization information systems is recommended as a strategy to improve the immunization coverage rate[10].

Northern Manhattan Immunization Partnership (NMIP)

The Northern Manhattan Immunization Partnership (NMIP)[11] is one of the four projects funded by the CDC to improve children immunization coverage rate through cooperation among academic medical centers, community health care providers and organizations seeking immunization rate improvement. The backbone of the NMIP is a computer-based multi-institution immunization registry[12]. Components of the registry include an immunization database, which is a relational model database implemented in Oracle7™ Release 7.3.3 (Oracle Corporation, Redwood City, CA), hospital registration systems, a Web-based registry server, WWW user interface and a decision support system.

The decision support system has multiple functions, including presenting on-screen reminders of vaccines due, checking allergies and contraindications, and generating recalls to health care providers and patient parents via email, fax or regular mail. The reminders, recalls and checks of allergies and contraindications are based on the CDC's recommendations.

In this paper, we will describe our work of translating the CDC's childhood immunization guidelines into a computer-based reminder recall system. The current version of our translation is based on CDC's recommended schedule for the period from January 1998 – December 1998[13].

Principle of Translation

To translate text-based clinical guidelines into computer-based algorithms, two issues need to be addressed, i.e., our requirements to the algorithm and the goal of the translation.

Other researchers' experiences show that successful computer-based guidelines should be simple and can make use of available data. The goal should be realistic and reasonable[14]. Based on these principles, we decided that the reminder recall system should be a supporting function of the registry to assist health care providers to make decisions instead of making decisions on their behalf. The system should be based on the existing New York Presbyterian Hospital's decision support system[15,16], the knowledge base of which was implemented as MLMs encoded in Arden Syntax[17,18].

As the first step of this work, we implemented the reminder recall system for the six routinely administered vaccine series in children, i.e., hepatitis B (HepB) vaccine series, diphtheria-tetanus-pertussis (DTP) vaccine series, Haemophilus influenzae type b (Hib) vaccine series, polio vaccine series, measles-mumps-rubella (MMR) vaccine series and varicella vaccine series. For the first stage of the implementation we used tools embedded in the Oracle database.

Architecture

We adopted a hierarchical modular approach in the design of the decision support architecture, as shown in Figure 1.

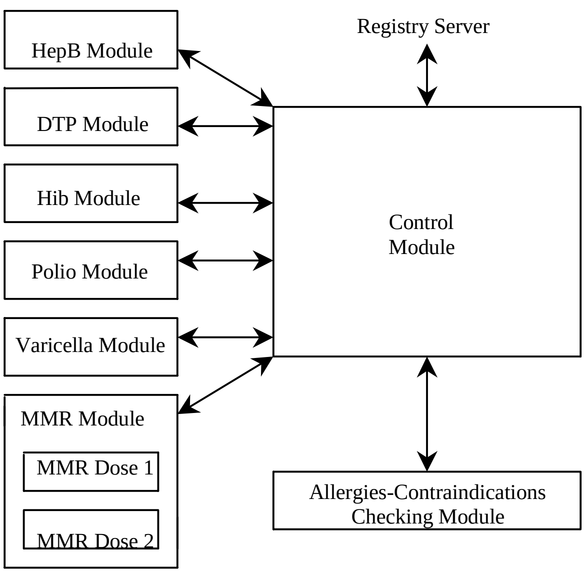

Figure 1. Hierarchical Modular Architecture

The *control module* is the central component of the system. It also serves as an interface to the registry server. When the reminder recall system is invoked by the registry server, the *control module* will check the possible allergies and contraindications by invoking the *allergies-contraindications checking module*. Based on this checking, appropriate vaccine series module will be invoked. Reminder recall messages will be generated, stored and feedback to the registry server.

Within vaccine series modules, the logic for each vaccine doses is encoded as second-level modules. For example, in MMR vaccine series, the logic for the two doses are encoded as two separate second level modules, as shown in Figure 1 (other second-level modules are omitted). Communication between these second-level modules is realized through their upper level module, i.e., the correspondent vaccine series module at the first level.

METHODS

Knowledge Representation

Based on other researchers' successful experiences[19], procedural and tabular approaches are combined to represent the immunization knowledge.

We use a tabular approach to represent knowledge that is easily expressed in a declarative way or needs to be updated regularly. For example, the list of vaccine series, the category of allergies and contraindications and the vocabulary codes are represented in tabular form. The vaccine dose number calculated in former invocations, which can be taken as knowledge in implicit form, is also represented in tabular form.

The logic for calculating the vaccine dose number and vaccine due date is naturally expressed by procedures. Procedural approach is used in representation of such logic for ease of understanding and similarity to human reasoning.

Language

In the first stage of the project, we tried to focus on the translation of the guideline logic rather than building a complex system with full features. Ease of implementation is the major consideration. We selected PL/SQL™ (Oracle Corporation, Redwood City, CA) to be the implementing language at this stage[20].

PL/SQL is provided by Oracle as a procedural extension of SQL. As a superset of SQL, it has all the features of SQL, which means it can provide easy and efficient operations on data in the Oracle database. Meanwhile, PL/SQL has a modular structure and procedural syntax. Procedural knowledge can be naturally encoded with branches and loops. This makes future implantation easy. Another advantage is that PL/SQL's transaction results can be cross-confirmed by SQL. In fact, in our informal evaluation of the translation, SQL is used to test the correctness of the program..

Avoiding Logical Incompleteness, Redundancy and Inconsistency

Logical incompleteness, redundancy and inconsistency are major considerations in translation of the guidelines.

With our procedural approach of knowledge representation, we use nested branches with hierarchical combinations of single antecedent variables to avoid logical incompleteness, redundancy and inconsistency. Here the antecedent variables refer to the variables in the left-hand side (LHS) of a procedural rule. For example, in the second dose of the hepatitis B vaccine series, antecedent variables include age when the first dose of hepatitis B vaccine is administered (for normal vaccine and catch-up vaccine we generate different messages), current age and time intervals since the first dose is administered. We nested these three variables hierarchically with a single variable LHS at each level to solve the incompleteness, redundancy and inconsistency problem, as shown by the fragment of logic for the second dose of the hepatitis B vaccine in Figure 2.

In immunization setting, many antecedent variables are not independent, which may result in tautological or invalid branches. For tautological branches, LHSs at some levels will always be satisfied when LHSs at other levels are satisfied. For invalid branches, LHSs at some levels will never be satisfied when LHSs at other levels are satisfied. Tautological and invalid branches are complimentary, i.e., if there are some tautological branches, there must be some invalid branches elsewhere, and vice versa. Although they are not errors, the tautological branches and invalid branches will make program inefficient and unnecessarily lengthy.

To address this problem, the co-relationship of the antecedent variables is analyzed. Those branches with invalid combination of antecedent variables are deleted to make the program efficient and concise. We use the second dose of the hepatitis B vaccine series as an example to illustrate it, as shown in Figure 2. The three antecedent variables, i.e., age when the first dose of hepatitis B vaccine is administered, current age and time interval since the first dose is administered, are not independent. From any two of them we can get the rest one in any specific combination of these antecedent variables. For example, if a child received the first hepatitis B vaccine dose at age of less than 3 months and his/her current age is less than 1 month, we can conclude that the time interval since the first dose must be less than 1 month. The branch with antecedent variable combination of the first hepatitis B vaccine dose administered at age of less than 3 months, current age of less than 1 month, and the time interval since the first dose equal or greater than 1 months is invalid, so we can delete this branch. The invalid branches are in bold as shown in Figure 2.

```
IF (age when dose1 is administered < 3 months)
THEN
    IF (current age < 1 month)
    THEN
        IF (time interval since dose1 is administered < 1 month)
        THEN
            (wait until due time);
        ELSE
            (invalid branch);
        END IF;
    ELSE IF (current age >= 1 month) AND
              (current age < 5 months)
    THEN
        IF (time interval since dose1 is administered < 1 month)
        THEN
            (wait until due time);
        ELSE
            (vaccine due);
        END IF;
    ELSE IF (current age >= 5 month) AND
              (current age < 11 years)
    THEN
        /* omitted */
    ELSE IF (current age >= 11 years) AND
              (current age < 13 years)
    THEN
        /* omitted */
    ELSE IF (current age >= 13 years)
    THEN
        /* omitted */
    END IF; END IF; END IF; END IF; END IF;
ELSE
    /* omitted */
END IF;
```

Figure 2. Nested Branches

Mapping to Vocabulary

To facilitate integration with the New York Presbyterian Hospital’s decision support system, concepts in the immunization reminder recall systems are mapped to the New York Presbyterian Hospital’s electronic medical vocabulary, the Medical Entity Dictionary (MED)[21,22]. With this mapping, and with our future work of implementation of the system in MLMs encoding in Arden Syntax, the immunization reminders recall system will be easy to be incorporated into the New York Presbyterian Hospital’s integrated information system architecture[23,24].

RESULTS

We implemented six vaccine series modules, one control module and one allergies-contraindications checking module. Inside the six vaccine series, 26 second-level vaccine dose modules with 195 original branches and 121 final branches after pruning the invalid ones are encoded. The number of second-level modules, original branches and final branches after pruning in each vaccine series are shown in Table 1.

	HepB	DTP	Hib	Polio	MMR	Var
2^{nd}-level modules	3	7	4	8	2	2
original branches	64	38	29	47	10	7
final branches	27	28	16	35	8	7

Table 1. Second-Level Modules and Branches

In the first load of data, 26311 immunization records of 5651 patients in 14771 visits were processed by the system. 41152 reminders (including reminders of prospective vaccines and notices of vaccines already past due) were generated. Manual queries to the database with SQL are performed to test the correctness of the reminders. 136 (0.33%) reminders related to 17 patients are found to be incorrect. Review of the data reveals they are due to the errors in the reload process of the data. There are 199 instances of data in which the birth date of the patient is after the vaccine date. In these instances, there are 6 patients whose birth date is in the future. Manual review found the year of their birth dates 19xx are erroneously recorded as 20xx. The rest of the instances have a year of their vaccine dates as 0097, which should be 1997. The errors are due to the bugs in conversion of dates between different systems when the data were reloaded from the New York Presbyterian Hospital's mainframe database.

DISCUSSION

Architecture

An advantage of our modular architecture design is ease of maintenance for both knowledge and programming. To incorporate new features, we can add new modules into the system. To modify the logic, for example, when updated version of the guidelines is published (as we mentioned before, the CDC updates their immunization guidelines regularly), we only need to update the specific modules involved. During this process, the modules not directly affected by the integration or modification can be kept intact. The current version of our translation is the second version. In our modification experience of the translation from the old version to the current version (due to CDC's guidelines modification), the modular design makes our work much easier.

An object-oriented approach has been used for guideline architecture design in other research efforts[25]. But up to now the only standard knowledge sharing mechanism, the Arden Syntax, is still in modular structure[26].

Knowledge Representation

Tabular form of knowledge representation makes maintenance easy. Other research efforts in immunization knowledge representation go a further step to store immunization parameters, such as minimum age for vaccine series and minimum time intervals between doses, in tables[19]. However, a tabular representation depends on table schema and its semantics, which is decided by the underlying logic. Tabular representation can facilitate maintenance under the assumption that the underlying logic is fixed and only the parameters are changed. How to facilitate the maintenance when the underlying logic is also changed need to be studied.

Procedural representation of immunization logic is easy to understand. Procedural approach in calculation of vaccine dose number and vaccine due date is similar to human reasoning. This will also make future work easy to encode the logic in Arden Syntax, which has a procedural syntax.

Logical Incompleteness, Redundancy and Inconsistency

Nested branches with hierarchical combinations of single antecedent variable are used in our implementation to avoid logical incompleteness, redundancy and inconsistency.

In general, if we have N antecedent variables V_1, V_2, …, V_N, and there are C_1, C_2, …, C_N categories for V_1, V_2, …, V_N separately, we need $C_1 \cdot C_2 \cdot \ldots \cdot C_N$ nested branches with hierarchical combination of single antecedent variables to avoid logical incompleteness, redundancy and inconsistency.

In fact, this approach is a variant form of truth table with explanations for antecedent variables. In immunization setting, the number of antecedent variables and their categories are relatively small. For complex problem, this approach may cause explosion of branches.

FUTURE WORK

A successful decision support system needs to be formally evaluated. The effectiveness of the reminders and recalls to the guideline compliance will be studied in our future work. The system will be encoded into MLMs using Arden Syntax, and we will study this translation process.

Acknowledgement

This work was supported by grant U66/CCU212961-03 from the United States Centers for Disease Control and Prevention.

References

1. Centers for Disease Control and Prevention. Vaccination coverage of 2-year-old children – United States, 1993. MMWR Morb Mortal Wkly Rpt 1994;43:705-9.
2. Centers for Disease Control and Prevention. State and national vaccination coverage levels among children aged 19-25 months – United States, April-December 1994. MMWR Morb Mortal Wkly Rpt 1995;44:613,619-23.
3. Wood D, Donald-Sherbourne C, Halfon N, et al. Factors related to immunization status among inner-city Latino and African-American preschoolers. Pediatrics 1995;96:295-301.
4. Strobino D, Keane V, Holt E, Hughart N, Guyer B. Parental attitudes do not explain underimmunization. Pediatrics 1996;98:1076-83.
5. Available at: http://aepo-xdv-www.epo.cdc.gov/wonder/PrevGuid/PrevGuid.htm. Accessed Feb.26, 1999.
6. Lieu TA, Black SB, Sorel ME, Ray P, Shinefield HR. Would better adherence to guidelines improve childhood immunization rates? Pediatrics 1996;98:1062-8.
7. Centers for Disease Control and Prevention. General Recommendations on immunization recommendations of the Advisory Committee on Immunization Practices (ACIP). MMWR Morb Mortal Wkly Rpt 1994;43(RR01):1-38.
8. Lobach DF, Hammond WE. Development and evaluation of a Computer-Assisted Management Protocol (CAMP): Improved compliance with care guidelines for diabetes mellitus. Proc Annu Symp Comput Appl Med Care 1994:787-91.
9. Overhage JM, Tierney WM, Zhou X, McDonald CJ. A randomized trial of “corollary orders” to prevent errors of omission. J Am Med Inform Assoc 1997;4:364-75.
10. Sinn JS, Kronenburg MA, Morrow AL. The purpose and functions of immunization information systems within health care organizations. Arch Pediatr Adolesc Med 1997;151:615-20.
11. NMIP Web site. Available at: http://nmip.cpmc.columbia.edu/. Accessed Feb.26, 1999.
12. Jenders RA, Dasgupta B, Mercedes D, Clayton PD. Design and implementation of a multi-institution immunization registry. Medinfo 1998:45-9.
13. Available at: http://www.cdc.gov/epo/mmwr/preview/mmwrhtml/00053300.htm. Accessed Feb.26, 1999.
14. Tierney WM, Overhage JM, McDonald CJ. Computerizing guidelines: factors for success. Proc AMIA Annu Fall Symp 1996:459-62.
15. Hripcsak G, Clayton PD, Jenders RA, Cimino JJ, Johnson SB. Design of a clinical event monitor. Comput Biomed Res 1996;29:194-221.
16. Jenders RA, Hripcsak G, Sideli RV, et al. Medical decision support: Experience with implementing the Arden Syntax at the Columbia Presbyterian Medical Center. Proc Annu Symp Comput Appl Med Care 1995:169-73.
17. Hripcsak G. Writing Arden Syntax medical logic modules. Computers in Biology & Medicine 1994;24:331-63.
18. Hripcsak G, Ludemann P, Pryor TA, Wigertz OB, Clayton PD. Rationale for the Arden Syntax. Comput Biomed Res 1994;27:291-324.
19. Miller PL, Frawley SJ, Sayward FG, Yasnoff WA, Duncan L, Fleming DW. Combining tabular, rule-based, and procedural knowledge in computer-based guidelines for childhood immunization. Comput Biomed Res 1997;30:211-31.
20. Urman S. Oracle PL/SQL programming. Berkeley: Osborne McGraw-Hill; 1996.
21. Cimino JJ, Hripcsak G, Johnson SB, Clayton PD. Designing an introspective, controlled medical vocabulary. Proc Annu Symp Comput Appl Med Care 1989:513-8.
22. Cimino JJ, Clayton PD, Hripcsak G, Johnson SB. Knowledge-based approaches to the maintenance of a large controlled medical terminology. J Am Med Inform Assoc 1994;1:35-50.
23. Hendrickson G, Anderson RK, Clayton PD, et al. The integrated academic information management system at Columbia Presbyterian Medical Center. MD Computing 1992;11:35-42.
24. Clayton PD, Sideli RV, Sengupta S. Open architecture and integrated information at Columbia-Presbyterian Medical Center: The role of networks, workstations and databases. MD Computing 1992;11:297-303.
25. Ohno-Machado L, Gennari JH, Murphy SN, et al. The Guideline Interchange Format: a model for representing guidelines. J Am Med Inform Assoc 1998;5:357-72.
26. American Society for Testing and Materials. E1460 standard specification for defining and sharing modular health knowledge bases (Arden Syntax for Medical Logic Modules). ASTM Standards, v14.01. 1992;539-87.