# Design Patterns for Clinical Guidelines

**Mor Peleg, PhD[a,b] and Samson W. Tu, MSc[a]**

[a]*Stanford Center for Biomedical Informatics Research, Stanford University, Stanford, CA, USA, 94305*

[b]*Department of Management Information Systems, University of Haifa, Haifa, Israel, 31905*

**Communication concerning the manuscript and reprint requests should be addressed to:**

Mor Peleg

Phone: +1 (408) 733-1531

Fax: 1-413-375-3755

Email: peleg.mor@gmail.com

**Present address:**

10367 Vista Knoll Blvd.

Cupertino, CA, 95014, USA

## Summary

**Objective**: Transforming narrative guidelines into a computer-interpretable formalism is still a bottle-neck in the development of decision-support systems. Our goal was to support this step by providing computer-interpretable templates for representing guideline knowledge using clinical abstractions that are appropriate for particular guideline sub-domains.

**Methods and materials**: We analyzed guidelines taken from the sub-domains of screening and immunization guidelines to find repeatable clinical abstractions and structured them as design templates to support encoding of these guidelines in a computer-interpretable format. To find guidelines for analysis and validation, we (1) searched the National Guideline Clearinghouse for screening guidelines in internal medicine, that have a clinical algorithm, and which were published during 2002-5 and (2) used adult and childhood immunization guidelines developed by Center of Disease Control and Prevention (CDC) and the Institute for Clinical Systems Improvement.

**Results**: We developed two visual templates that structure screening guidelines as algorithms of guideline steps used for screening and data collection and used them to represent the guidelines collected. We validated the computability of the screening templates by executing a screening guideline in a workflow engine. We validated the computability of immunization templates by writing code that, based on represented knowledge, computes immunization due dates and by creating an algorithm that translates the knowledge into computer-interpretable guidelines.

**Conclusion**: We have demonstrated that our templates could be effectively applied to screening and immunization guidelines to produce computer-interpretable representations using domain-level abstractions.


**Key words**: clinical guidelines, design patterns, immunization guidelines, screening guidelines, guideline modeling

## 1. Introduction

Clinical guidelines are "systematically-developed statements to assist practitioner and patient decision making about appropriate healthcare for specific clinical circumstances" [1]. They aim to improve the quality of patient care, minimize unjustified variance in patient care, and cut down costs. Studies have shown [2] that computerizing guidelines, implemented as decision-support systems (DSS) that provide patient-specific advice at the point of care, have better impact on clinicians' behavior as compared to narrative guidelines, hence such implementations are more likely to achieve the goals of clinical guidelines. Therefore, many researchers have been developing languages to represent guidelines in computer-interpretable formats [3, 4]. Yet, the amount of work and effort involved in computerizing narrative guidelines is immense: narrative guidelines are often ambiguous, inconsistent, and incomplete [5] and the narrative guidelines need to be re-conceptualized in terms of the modeling constructs of computable guideline formalisms [6]. Therefore, when teams of information scientists and clinicians computerize such guidelines, they cannot simply translate the text and tables into a computable format; the recommendations of guidelines needs to be operationalized in terms of available data, hidden assumptions need to be made explicit, and the undefined terms and inconsistent recommendations need to be clarified by the teams. Such intrinsic shortcomings of narrative guidelines constitute one of the major barriers toward successful guideline implementations [7].

In order to ease the encoding of narrative guidelines in a computer-interpretable formalism, tools have been developed to support the markup of guideline text in electronic format [8-10] to ease its transformation and representation in a computable guideline formalism [9, 10], and even to validate properties of the encoded representation [11, 12]. However, the transition into an encoded guideline requires formulating the marked-up guideline using low-level guideline components such as clinical tasks, actions, decisions, and criteria that are foreign to a clinician's cognitive process.

In this paper, we propose that domain-specific high-level computer-interpretable *design patterns* can significantly ease the process of encoding computer-interpretable guidelines. In Section 2, we describe the formulation of such design patterns for screening and immunization guidelines as

templates that can be instantiated for specific guidelines. In Section 3, we demonstrate the computability of the instantiated templates by using them as inputs into the Bonita workflow management system (http://wiki.bonita.objectweb.org/xwiki/bin/view/Main/WebHome) and the Jess rule system (http://herzberg.ca.sandia.gov/). In the rest of this section, we discuss the repeatable patterns that other research groups have already found in clinical guidelines [10, 13-15] and the available representation of immunization and screening knowledge, and conclude the section by outlining our approach and its relationships to these efforts.

### 1.1. Design patterns in clinical guidelines

A design pattern [16] is a general repeatable solution to a commonly occurring problem in system analysis and design. It is a description or template specifying how to solve a problem that can be used in many different situations. Design patterns can ease conceptualization of a problem by a system analyst, facilitate transformation into code, promote reuse of problem specification and its coded solution, and provide a recommended specification that would achieve some goal (e.g., a DSS design that minimizes unnecessary interaction with users).

Researchers from the medical informatics community have developed different guideline modeling templates. Their approaches differ in their purpose, scope, and characteristics. Following is a discussion of these approaches. Table 1 summarizes the main characteristics of these guideline template approaches.

Approaches developed in the Protocure-II project [14, 17] focused on using design templates to ease the transformation from narrative guidelines into encoded formalizations. Their approach addressed the translation of single recommendations into a formal representation by relying on clinically-based as well as data-type-based patterns without considering visualization of patterns. Developed by the Protocure-II project, the Many-Headed Bridge (MHB) [17] is an intermediate representation between informal (free text and tables) guideline representations and more formal guideline representations, such as Asbru [18]. The Document Exploration and Linking Tool (Delt/A) [10, 17] tool supports the transformation from a HTML (or a XML) document to any XML language using macros (i.e., design

patterns or templates) specified for the particular XML language. Using MHB, Delt/A supports a guideline modeler in manually transforming clinical guidelines from their original textual form (HTML) first to an intermediate and a semi-formal representation (MHB in XML) and finally to a formal representation in the Asbru language [18]. It contains macros for MHB for the following dimensions: control flow, data flow, temporal aspects, evidence, background information, resources, patient related aspects, and document structure.

In the Protocure-II project, linguistic patterns have been developed, which can be used to formally represent the knowledge about medical actions contained in text [14] . For example, the text "Patients with locoregionally advanced breast cancer should receive multidisciplinary treatment with curative intent" can be tagged and structured according to a recommendation pattern: med_context ("Patients with locoregionally advanced breast cancer") followed by a recommendation_op ("should") and an instance of med_action ("receive multidisciplinary treatment with curative intent"). These patterns have been used to automatically parse recommendations according to the defined action-based patterns.

To support guideline implementers in standardizing and implementing the action components of guideline recommendations, Essaihi, Michel, and Shiffman defined the Action Palette [15] – a set of medical action types that categorizes activities recommended by clinical guidelines. Guideline implementers might select action types from a palette-like implementation tool—much as artists select colors from theirs—to “paint” implementation activities in a standardized manner. The set of actions include: Prescribe, Perform therapeutic procedure, Educate/Counsel, Test, Dispose, Refer/Consult, Conclude, Document, Monitor, Advocate, and Prepare. The intention is to develop commonly used services for each action type, thus easing workflow integration.

An approach that was developed within the SAGE [19] guideline formalism is the use of expression templates to ease the formalization of clinical expressions. These expression templates allow encoders to easily author computable expressions without concern for expression language syntax. The templates are based on information models of clinical statements, use objects to represent the structure of expressions and include stereotypical eligibility and decision criteria, variables with derivation

expressions, functions that implement a very limited subset of the GELLO expression language [20], and queries on patient data and drug database based on specified information models.

All of the previously mentioned approaches addressed encoding of individual recommendation or parts of a recommendation and have not focused on visualization of patterns. Unlike these approaches, Boxwala et al. [13, 21] developed a visual risk-assessment Macro for the GLIF3 guideline formalism. Macros are sets of subclasses that can be used to represent patterns of domain-level concepts, aiding the guideline modeler in conceptualization of a clinical guideline or part of a guideline that may include more than one recommendation. Macros include attributes that define information needed to instantiate a set of underlying, non-domain-specific GLIF3 steps. For example, the risk-assessment macro includes attributes for collecting patient data, computing risk, and specifying recommendations based on risk. Currently, the mapping of macros to underlying GLIF3 steps is done manually.

### 1.2. Specification and representation of immunization and screening knowledge

While numerous studies report the use of decision aids in preventive services, few explicitly discuss their knowledge representation. We briefly survey some of these studies. IMM /Serve [22] is a computer-based guideline that produces recommendations for childhood immunization. It uses tables to store many of its temporal parameters. For each vaccine and dose, these parameters include the child's age and minimum waiting time from previous doses for different clinical conditions. If-then-else rules are used to model the logic that decides which set of parameters applies to a particular patient at a particular time and which vaccine preparation to use. Procedural code (in the C language) is used to translate different units of measure (of time), to schedule immunizations on a calendar, and to coordinate the system's operation. Finally, a decision table is used to store the screening parameters for each dose together with a small amount of clinical logic to determine whether previous doses are valid [23]. The developers of this system used domain knowledge about the temporal relationships and temporal constraints involved in childhood immunization to identify possible errors and contradictions within the knowledge base. They also developed a program that generates test cases that could be evaluated by clinicians. The program uses a rule-base to determine the next dates of

immunizations for individual cases. The preprocessing step takes a case description as input and generates a set of facts that can be checked by the rule engine. The rule interpreter then determines which set of age and interval constraints apply to the case. A post-processing step converts the applicable recommendations into dates.

Wang, Jenders, and Dasgupta [24] used a combined approach of tabular and procedural knowledge representation to develop the EzVac system that produces reminders for childhood immunizations. As in the IMM /Serve system, tables are used to store basic parameters, such as the list of vaccine series, the category of allergies and contraindications, the vocabulary codes, and dose numbers. While the developers of the IMM/Serve system tried to represent knowledge declaratively as much as possible, using decision tables, and resorted to if-then-else rules and to code only in complex tasks, the developers of EzVac used only procedural representation of immunization logic for calculation of vaccine dose number and vaccine due date. The procedural representation was written in Oracle PL/SQL or as medical logic modules (MLMs) written in the Arden Syntax.

The Child Health Improvement through Computer Automation system [25] uses MLMs to decide which pre-screener forms (i.e., a form that collects information from the patient or the parent before the physician encounter) are relevant to a child, based on criteria relating to age, and encoded as part of the MLMs. Vendors, such as Micromedex, are using Arden Syntax and MLMs to implement screening guidelines, including immunization screening (i.e., record review to determine which immunizations might be appropriate) [26]. Taking a process-based approach, screening guidelines have also been developed as computer-interpretable guidelines, encoded in clinical-guideline formalisms such as GLIF [27] and PRO*forma* [28] or in generic process models [29] that incorporate clinical domain ontologies. The Unified Modeling language has been used to formulate requirements, a data model, and a prototype DSS for an obesity screening and management guideline [30], using an approach that views the guideline as a process of care that unfolds over time. As these works show, screening guidelines have been implemented either using individual medical decision rules such as the Arden Syntax and MLMs, or using a process-based approach that views the guideline as a process of care that unfolds over time. As we will show in this paper (see figures 6 and 7), screening can be a

complex process. It thus may benefit from a process-based approach that explicitly shows how individual decision-making rules interact with each other such that the overall effect of modifying one part of the guideline could be more easily comprehended [31].

### 1.3 The approach used in this paper

Except for the GLIF Risk Assessment Macro, the other patterns mentioned in Section 1.1 focused on the level of individual recommendations or their component parts. That level of representation does not help with conceptualization of guidelines. Because these patterns focus on individual recommendations, templates derived from them do not help guideline users in seeing the "big picture" of the guideline as they encode guidelines. Moreover, these patterns do not support computer-based execution. Based on our experience in modeling clinical guidelines, we hypothesized that classes of guidelines have stereotypical recurring patterns for organizing and conceptually representing guideline knowledge in a computer-interpretable way. These patterns could span more than a single recommendation: aggregations of recommendations (guideline parts), or even entire guidelines, which describe a clinical procedure with a single clinical intent (e.g., screening, prevention, chronic disease management). Such patterns allow us to define guidelines that are closer to the understanding of the content expert and that are still computer-interpretable. Once defined, such templates could ease conceptualization and understanding by non-expert modelers, standardize their specification, enable more rapid encoding and easier maintenance, ease transformation into code, and facilitate reuse. This would enable a new division of labor, where experienced modelers could analyze and formulate patterns, and non-expert modelers would use these patterns in a standard way. Moreover, once such high-level patterns are defined, existing guideline template approaches could be combined with them to ease modeling of details.

In this paper we test our hypothesis by formulating models of design patterns for screening guidelines (or parts of guidelines that address screening) and for immunization guidelines and by demonstrating that such models lead to computer-interpretable encodings of these guidelines. We have chosen these domains because they illustrate two very different types of guideline knowledge, one that is

procedurally oriented while the other is constraint-based. Another reason for choosing these domains is that we leveraged from prior work that we did in the GLIF [32] and SAGE [33] projects and from the work of Miller, Frawley, and co-authors on a DSS for immunization guidelines [23]. Screening and immunization guidelines are very important and common, yet hold a significant amount of medical knowledge. Thus, it is extremely useful to help the conceptualization and standardization of clear, complete and sound guidelines of this sort. We also recognized recurring patterns in other classes of guidelines, for example in chronic-disease management guidelines. However, we have yet to formalize them as templates. In the Discussion section, we consider evidences for such stereotypical patterns in other domains and how other groups [34] have already identified modeling paradigms for them.

Unlike guideline modeling languages [4, 35], which offer generic constructs for representing any careflow process as a network of tasks conveying medical actions and decisions, the patterns that we have been developing address modeling of clinical knowledge in specific sub-types of guidelines. Our patterns offer templates that can be used to structure the clinical knowledge in a top-down manner that breaks the knowledge into manageable interacting components that are meaningful for the particular sub-types of guidelines. Thus, they could potentially relieve the so-called "analysis paralysis" [36], where there are few clues to guide the system analyst (or encoder in our case) in generating the guideline model from the generic non-clinical constructs offered by guideline modeling languages.

In the next section, we illustrate our modeling approach by focusing on two types of clinical patterns: process-oriented visual algorithms for screening guidelines and tabular immunization temporal constraints patterns. In Section 3, we validate our model by showing examples of guidelines encoding using these two types of clinical patterns, and demonstrate how they could be translated automatically into inputs for Bonita and Jess.

## 2. Methods

For each of the two patterns─process-oriented screening algorithms and temporal immunization constraint tables─ we describe the materials and process we used to formulate the patterns and describe in detail the templates that can be used to model such patterns.

## 2.1. Methods for screening algorithm patterns

### 2.1.1 Materials

We started by searching for guidelines from the National Guidelines Clearinghouse (www.guidelines.com) repository that were categorized as having the category of *screening* and clinical specialty of *internal medicine*, that had a clinical algorithm, and that were uploaded from the years 2002-5. After inspecting the 56 retrieved guidelines, we found out that only 9 of them were in fact guidelines for on-going screening, rather than one-time diagnosis or disease management. These guidelines included screening for colorectal cancer, cervical cancer, lipids, obesity, hypertension, and pediatric eye evaluations. We arbitrarily selected 4 of these 9 guidelines (lipid screening, obesity, cervical cancer, colon cancer) for developing our design templates and used the rest to verify the generality of the templates.

During the development of the screening algorithm pattern we used the knowledge modeling tool Protégé [37]. This tool allowed us to define the abstraction patterns as an ontology of screening algorithms, by defining classes of screening algorithm steps, their properties, and relationships. An ontology is an appropriate representation of knowledge, as it is an explicit specification of a conceptualization [38], which can be used to share common understanding of the structure of the information among people or software agents, to enable reuse of domain knowledge, and to support reasoning and computation. Our Protégé screening patterns include visual specification of a screening algorithm with another layer of computable definitions of parameters and decision points, supporting execution of the knowledge. Moreover, as we will demonstrate below, the process patterns that we developed may be specified with the Workflow formalism, which is more generic than our Protégé templates.

### 2.1.2 Identification of patterns in screening guidelines

During pattern development, we identified common themes in the four selected guidelines, which we structured as steps of the screening process. These common themes included a step that determines the time for the next screening test, a step for collecting all the data necessary in order to assess the patient, and a categorization step, which categorizes the patients into populations of patients who should continue screening and those who already screened positive for the disease. Then, we examined each of these four guidelines one by one, and tried to align them with the common steps. We looked for information that was contained in the narrative guidelines that was not captured by the existing screening pattern and extended the pattern so that it would capture it. For example, we found that some of the guidelines included additional steps (e.g., patient education in the obesity and colon cancer guidelines) that did not appear in all of the screening guidelines. Nevertheless, if such steps appeared in more than one guideline and were considered as important steps, we added them to the template.

After structuring the high-level view of a screening guideline, we addressed the more detailed knowledge of collecting the results of screening tests. Since for some screening guidelines (e.g., colon cancer screening), many alternative tests were available and detailed recommendations were provided for deciding whether to repeat a test or resort to alternative tests in case of inconclusive results, we studied these guideline in order to generalize recurring themes for obtaining screening results, which re-occurred in the alternative screening tests. In addition, some screening guidelines (e.g., the cervical cancer guideline) contained very detailed knowledge regarding the timing at which screening cycles should be done, as determined by results of previous screening results. We constructed the template by studying these guidelines. We refined the templates until all the information contained in all four guidelines could be represented conveniently using the templates that we had created.

After completing the template development, we inspected the other screening guidelines to see that they would fit the patterns. To do this, we read the guidelines and mapped parts of them to the different screening template steps, making sure that all guideline knowledge could be mapped to the various steps. We also examined the criteria for test selection and the recommendations about the timing of screening tests, and saw that they were similar to constraints that we had already created for

the first four guidelines. After finalizing the screening pattern, we used it to model two additional screening portions of guidelines that we have previously modeled without the screening pattern: diabetic foot [39] and retinopathy [40]. The templates accommodated the modeling of these test guidelines without the need to make modifications to the templates.

2.1.3. Screening templates

The screening patterns that we developed contain templates for screening and for data collection.

A) Screening Template

The template for a screening algorithm includes the following parts:

**Screening initiation condition**. For example, cervical cancer screening should be initiated for (women who are 21 years of age) or (women whose onset of sexual activity was over 3 years ago) or (women over 65 who have a new sexual partner)

**A condition for terminating screening**. For example, screening for cervical cancer should be terminated for (women who are over 65 who had 3 consecutive normal cervical cancer screenings in the last 10 years) or (women who had a hysterectomy for benign disease).

**Time-for-screening node** – screening is a procedure done for detecting disease or disorder in a defined population that does not have symptoms or signs of that disorder. Our screening templates were created to be generic enough to represent different types of screening procedures – ones that are done once per life time (e.g., genetic test) and ones that are done periodically (e.g., screening for cervical cancer). The Time-for-screening step is used for determining whether it is time to screen, based on the screening frequency set during the previous screening cycle and the time that has passed since the previous screening. This step uses a table that specifies the possible screening test that was done during the previous screening (specified as data item holding screening test result) and screening frequencies for certain patient conditions. When a screening test is to be applied only once in a life time, this construct would not be used.

**Performing-screening node**, which specifies things that should be done during screening. It is further decomposed into the Data Collection template discussed below. This step acquires all data needed for categorizing patients.

**Categorization node** for dividing patients into risk categories. This step has an optional slot for specifying preconditions that must be met prior to categorization (e.g., availability of test results).

**Conditioned screening connector** – used to connect the categorization step to one of the risk categories. This connector includes a criterion that refers to collected patient data item values. For example, in cervical cancer, screening continues for patients whose physical cervix exam and PAP smear were negative.

**Risk category node**– there are two types of categories. The Continue Screening category contains tables for determining screening frequency based on screening test results and other patient parameters. These tables have the same structure as the time-for-screening-tables. The second risk category type specifies categories of patients who test positive for the problem being screened.

**Patient Education node** – can be used anywhere within the algorithm.

B) Data Collection Template

The second template –Data Collection–is used to specify the tests that should be done to gather all necessary data so that the patient could be categorized. We found that in some screening guidelines, such as the colon cancer guideline, tests are repeated to confirm positive or negative results, or when the test results are inconclusive (i.e., the test was done well but the quality of its results prevented from establishing conclusive results) or inadequate (i.e., the test was not done well, for example, the colonoscopy apparatus could not reach the entire colon). Therefore, in order to collect all data necessary to establish a patient category ("having the disease" or "continue with screening" at the specified screening interval) the data collection pattern includes repeating tests when the results are inadequate, inconclusive, or need to be confirmed. We also found that in many cases, different alternative tests could be done in order to collect the same information about the patient. Accordingly, the Data Collection pattern defines an algorithm consisting of a network that includes three types of

nodes: test selection, test performance, and data collection. Conditions on connectors may specify conditional traversal of a network. For example, HPV_DNA is a test that needs to be performed only for women who are 30 years or older. Other conditions, such as "results available", "results inadequate", or "results inconclusive" may necessitate further test selection and test performance after the initial data-collection step. We recognize that the selection of appropriate screening test may involve complex considerations such as sensitivity, specificity, risks of adverse events, cost, patient preference, age, comorbidities, and the temporal patterns of benefit relative to harm [41], Employment of quantitative methods such as decision analysis and preference elicitation to mode the test selection problem is beyond the scope of the templates we use for representing guideline information. Nevertheless, our templates are still applicable to many guidelines where the guideline authors compiled the knowledge about the properties of individual screening tests (such as their specificity and sensitivity, and their risks) into decision criteria that are summarized as recommendations without the need for the guideline reader to get into such considerations for test selection.

C) Assumptions for screening patterns

In using the screening patterns, we made two simplifying assumptions.

Assumption 1: When you perform a test, you can assume that test results will become available sometime in the future. The model does not need to check whether some exception occurred and test results were not obtained in time.

Assumption 2: The screening tests (e.g., microalbumin/creatinine test, as explained below) were not done for other purposes. Therefore, we can assume that if there is a test result available (in the electronic medical record) then it represents the results from the last screening test. Without this assumption, the problem of recognizing screening due dates becomes that of *pattern recognition*. For example, in an diabetic nephropathy guideline that we encoded, an abnormal microalbumin/creatinine test results needs to be confirmed by one or more additional tests at least 3 months apart. If we try to avoid unnecessary tests by using results of tests not specifically ordered for screening for diabetic

nephropathy (e.g., tests done to detect albuminuria in hypertension and systemic lupus erythematosis), then we need to search for temporal patterns in patient data that allow us to confirm the presence of chronic kidney disease. Therefore, the pattern we describe in this paper is not appropriate for the diabetic nephropathy screening guideline.

D) A check list for screening patterns

Figure 1 shows the check list that we developed to help modelers avoid modeling errors. Our assumption was that the recommendations provided by the narrative clinical guidelines were clinically valid and that the guidelines are complete. Thus, we developed the check list in order to standardize the way in which the design patterns are used to model narrative guidelines and to prevent errors in modeling the knowledge found in the narrative guidelines. This check list is not intended to improve the accuracy and completeness of narrative guidelines or to allow readers to assess the potential for bias in test selection and reporting, as done, for example, in the Standards for Reporting of Diagnostic Accuracy (STARD) initiative [42].

## 2.2. Methods for immunization temporal constraints patterns

### 2.2.1 Materials

We formulated our model after studying the immunization plans (normal and catch-up schedules) for Pneumococcal vaccines PCV7 and PPV23, Polio virus and Diphtheria. We analyzed these published guidelines, finding common themes, as described in the next section. After developing the model, we also used it to specify the immunization schedules for Haemophilus influenzae B Conjugate (Hib), hepatitis-A, MMR/V, and Varicella, and inspected other immunization schedules to verify that the model is sufficient to encode all of the pediatric and adult immunization schedules published by Centers for Disease Control and Prevention in the U.S. [43].

### 2.2.2 Analysis of knowledge contained in immunization guidelines

Studying the published immunization guidelines, we found that they contain many complex criteria for determining the scheduling of immunization doses and catch-up schedules. We first tried to find the rationale for giving and for not giving immunization doses. We reasoned that, in principle, a patient should not be given an immunization if he is already immunized because he either (a) was given the total number of recommended doses (which may depend on current age and the age at which certain vaccinations were given), (b) had the disease, or (c) has antibodies for the disease. Then we examined reasons that were provided in clinical guidelines for now immunizing patients who are not immunized against the disease. Reasons not to immunize include:

- Contraindications
- Known allergy to a vaccine component or bad reactions occurring at a previous dose
- Absence of risk indications (i.e., indications for giving an immunization dose specifically to children at increased risk. For example, a fourth dose of PCV7 may be indicated because of certain chronic conditions such as chronic heart or lung disease).

Most of the recommendations concerned instructions regarding immunization schedules for patients who should be immunized. The immunization schedules were arranged in tables. We saw that these tables related the schedule to:

- The child's current age (date of birth, DOB)
- Previous SubstanceAdministration given (important data include immunization_type (e.g., Diphtheria), specific vaccine (e.g., DTaP), and date given)
- Recommended wait times between vaccinations
- Catch-up schedules, often provided in separate tables

Tables are a common and useful representation for immunization schedules in published guidelines. Yet, as shown in Figure 2, the information stored in the tables is incomplete and is supplemented by text and footnotes. That information described specific vaccine types, missed doses, invalid doses,

boosters, and other interacting immunization types. We wanted to represent all the information in a single table-based pattern.

2.2.3 Identification of patterns in published immunization guidelines and SAGE encodings

We examined the way in which the SAGE project [19] represented immunization knowledge used for providing decision-support for checking whether immunizations are due on certain dates. Figure 3 shows a screenshot from the SAGE guideline model for immunizations, specified using Protégé [37].

To encode this immunization knowledge, clinical experts wrote the criteria for giving a particular dose of a vaccine in non-computable, stylized English and then knowledge engineers encoded the criteria and verified with the same clinicians that the encoded criteria are consistent with the clinical knowledge. The criteria are encoded as instances of criterion classes, which include among others, Complex Criteria that include criteria connected with logical connectors of And/Or/Not (such as the two frames on the top of Figure 3), Presence/Absence Criteria (e.g., presence of asthma), Comparison Criteria (e.g., age <18), Temporal Comparison Criteria (e.g., time of first PCV7 vaccination > 28 days – shown on the bottom of Figure 3). The problem with this representation is that the encoded criteria are similar to the antecedents of discrete rules and it is not possible to check for completeness of or inconsistencies in the criteria sets; appropriate visualization of relationships between time-oriented knowledge is important for user comprehension [45] and can facilitate the discovery of inconsistencies manually. In addition, this representation did not use clinical patterns specific to immunization, but rather generic temporal-comparison templates.

We found the following patterns of information repeatedly appearing in immunization guidelines (published guidelines and SAGE encodings): (1) total number of doses, (2) minimum and maximum age requirements for particular doses, (3) the vaccine type and age at which some earlier dose is given, (4) spacing between doses, (5) risk indications, and (6) contraindications associated with doses.

The formulation process of the immunization patterns was an iterative process, in which we examined the different combinations in which the above six primitive patterns appeared in immunization patterns. We started with the Pneumococcal immunization guideline and, using the Protégé tool,

Comment [MP1]: R1.5

defined for it immunization-specific (not generic temporal comparisons as in SAGE) clinical templates that combine the six primitive patterns in certain ways. Our first formulation template assumed a separate template for each vaccine type (e.g., PCV7). The template specified for each vaccine type the total number of doses (pattern 1) and contraindications associated with doses (pattern 6). Restrictions for each dose were specified using a separate template. That template specified minimum and maximum age requirements (pattern 2), the vaccine type and age at which an earlier dose is given (pattern 3), spacing between doses (pattern 4), and risk indications (pattern 5). However, that template had several shortcomings. First, pattern 3 related only to previous doses of the same vaccine type. Second, the spacing between doses could relate to any vaccine type, not to just one vaccine type. But most importantly, that representation could not assist us in checking for completeness or consistency of restrictions; it was very difficult to tell for a given dose, which restrictions related to a given age requirements and whether all age requirements were specified.

Therefore, we moved to a tabular representation, described in the next subsection, in which completeness and consistency of restrictions was much more easily recognized; the tabular representation tabulated for each dose, the earlier dose constraints against the age intervals. For each combination of these two axes, the spacing between doses and the vaccine to be given were specified. We initially represented the Pneumococcal vaccines, followed by Haemophilus influenzae type b vaccine (Hib) and Hepatitis A vaccines. The tabular representation that we used initially did not address invalid doses (doses that were given too early after a previous immunization or at an age that was too young), inadvertent use (doses which were not the age-appropriate vaccine types), and the flexibility of deviating from the minimum wait time or age constraints. In addition, our early models did not address multiple earlier dose constraints that were all applicable to the same dose and age restrictions and they also did not address earlier dose constraints that related to more than one possible vaccine type (e.g., DTaP, Td, Tdap). When we encountered such complicated conditions in the Mumps/Measles/Rubella, Varicella, and Diphtheria immunization guidelines, we found that our tabular format could accommodate them with minor changes that did not affect the table structure but instead, by introduced new notations for the table's cells. These changes include the addition of a wild

card notation for specifying any vaccine out of a list of vaccine types, a notation for specifying doses that are not counted (NC) due to inadvertent use and replacement doses for them. Our final tabular representation is described in the next subsection. The algorithm that calculated the due date of the next vaccination was revised to handle these new notations and to calculate invalid doses and inadvertent use. It is described in Section 3.4.1.

### 2.2.4 Immunization template

As discussed in sections 2.2.2 and 2.2.3, tables found in published guidelines and guideline modeling languages' representation of criteria that determine the timing and type of vaccines to be given had a number of shortcomings. Therefore, our challenge was to represent in a single representation the constraints that combine information about the type and age at which past vaccine have been given, the current age of the child, the temporal constraints on the wait period from past vaccines, and the type of vaccine that should be given for each immunization type. Our immunization template models temporal constraints among vaccination doses in a tabular format, allows consistency checks during authoring, and is translatable to computable format, as described in the Results section. A design decision we made was that we would not to address the encoding of indication and contraindication constraints (patterns 5 and 6 described at the end of the previous subsection), as they are already well-supported by other formalisms, such as SAGE and the GELLO expression language [20]. Our immunization patterns could be easily extended with GELLO or SAGE's representation of indication and contraindications for each dose and vaccine. We further assume that a patient's immunization records are complete.

Our immunization template supports modeling of temporal constraints with several possible vaccines (e.g., DTaP, DT, and Tdap). This template allows the modeling of patterns 1-4 described at the end of the previous subsection. In section 3.3 we show, in Excel, the immunization schedule of diphtheria vaccine as an example of our immunization table specification.

**Pattern 1**- current dose number: the immunization template contains a tabular representation for each dose. For example, row 3 in Figure 4 specifies the third dose of the Diphtheria immunization.

The immunization template superimposes patterns 2, 3, and 4 (minimum and maximum current age for a dose, vaccine type, age ranges, and dose number of earlier dose, and spacing before dose) into a single table (one table per dose).

**Pattern 2 -** current age for a particular dose: the column headings of the immunization template shown in the right part of rows 5 and 6 in Figure 4 specify the ranges of the current age for a particular dose

**Pattern 3 -** earlier dose constraints: these include the vaccine type, dose number, and administration age of an earlier dose. They are represented as rows in the immunization template (see the left part of rows 5 and 6 in Figure 4

**Pattern 4 -** spacing between doses: The minimum waiting time since the previous dose and the vaccine that should be given are specified in the cells of the right part of the immunization template (e.g., see the text on the bottom of Figure 4:"Minimum waiting time since previous dose").

We model the format of the immunization template as a Protégé ontology. The ontology has the following classes, as shown in the UML class diagram of Figure 5:

***Immunization*** is a class holds an immunization plan for a certain immunization type (e.g., diphtheria). It enumerates possible vaccines for the immunization type and points to a set of Immunization Dose Constraints (one for each dose).

***Immunization Dose Constraints*** is a class whose instances that hold sets of constraints for a specific dose number, (i.e., instances of the Immunization Constraint Set class, described below), which are defined for the dose number pattern (pattern 1), and correspond to a table such as that shown in Figure 4.

***Immunization Constraint Set*** instances hold two constraint types: Current Age Constraint and Next Vaccination Constraint, described below;

***Current Age Constraints*** hold constraints regarding the current age that is appropriate for the dose number specified in *Immunization Dose Constraints* (i.e., those represented by pattern 2)

***Next Vaccination Constraint*** references an Earlier Dose Constraint (see below) and for it specifies the minimum interval from previous dose and the recommended vaccine (i.e., information found in the cells of the immunization template, corresponding to pattern 4).

***Earlier Dose Constraint*** corresponds to pattern 3 and specifies the type of earlier dose, its dose number, and the age interval during which it was given.

In Section 3.3 we show a screen shot of the diphtheria immunization specified using the Protégé ontology.

## 3. Results

We validated the screening and immunization temporal constraints patterns by using the templates to model the guidelines we selected for the study. As explained in the Methods section, the guidelines used for validation were different from those used for model formulation. In addition, we demonstrate in this section how we used knowledge modeled in the immunization pattern to drive automated decision support.

### 3.1. An example for the screening templates

Figure 6 shows a screen shot of the cervical cancer screening guideline, using the screening template in Protégé. Table 2 shows the time-for-screening table, corresponding to the first step in the screening template, taken from the cervical cancer model. The first row specifies that if the last test result is of PAP smear, and it was normal, the physical exam of the cervix was normal and HPV DNA test was positive, then the screening frequency should be every 6-12 months. The other parts of Figure 6 are explained in the legend of the figure. The Data Collection for Cervical Cancer Screening node expands into the Data Collection pattern.

Figure 7 shows a screen shot of the colon cancer Data Collection pattern in the workflow editor of Bonita (http://wiki.bonita.objectweb.org/xwiki/bin/view/Main/WebHome) – an open-source workflow

tool that we used for creating and executing workflows corresponding to the screening and data collection patterns. In this tool, the three logical types of steps in the template (Selecting Test, Performing Test and Obtaining data result) are not distinguishable by their appearance as they are in the Protégé tool. However, the names of the steps are indicative of their types. For example, activities named "Selecting Test" indicate test selection. Activities named "Inspecting Results" indicate obtaining a test result. The other nodes, with names of specific tests (e.g., FOBT) indicate performance of tests. Figure 7 shows how these steps are connected through conditional links to end at a state (End) where all data necessary to categorize the patient has been gathered.

### 3.2. Validation of the screening patterns' computability

We used the Bonita workflow engine to execute the colonoscopy guideline, entering user input regarding test selection and test results.

### 3.3. An example for the immunization temporal constraints template

Figures 4 and 8 show examples of the use of our immunization temporal constraint pattern to encode the immunization schedule for diphtheria vaccine. The examples implement, in Excel, the temporal constraints relevant to the administration of the third dose of diphtheria vaccine. We start with Figure 4, which represents a simplified version of the immunization constraint set to demonstrate the important principles of our approach. We then explain the full constraint set, shown in Figure 8.

Row 1 in Figure 4 shows the possible vaccines for diphtheria: DTaP, Td, Tdap. *Relevant prior vaccines*, shown on Row 2 (DTaP, Td, Tdap), are earlier vaccinations that may affect the timing and type of the next diphtheria vaccine. Row 3 specifies the *dose number* – which is 3 in our case. Row 4 specifies the *recommended age for the dose*, which is 6 months for the third dose. Below it is the main specification table. The right part of the table has columns for *ranges of the subject's current age*: 0-14 weeks, 14 weeks-7 years, 7-11 years, and 11-18 years. The left part of the table shows details of earlier doses that may affect the timing and type of the current dose. These details include the *type* of the earlier dose (where * is defined at the upper part of the table to mean any relevant prior dose, i.e., DTaP, Td, or Tdap), the dose number of the earlier dose and the age range (minimum time, maximum

time) during which the earlier dose was given. Each *row* lists a different *earlier dose constraint*. For example, the earlier-dose constraints (left) part of Row 7 specifies the timing and type constraints imposed by the first dose of diphtheria (DTaP, Td, or Tdap) that was given when the child was of age between zero and 1 year. Each *cell* in the right part of the table corresponds to a certain earlier dose constraint and to the current age of the child. Each cell specifies the *minimum wait period* and below it the vaccine that should be given after the wait period, counted from the last vaccination of the immunization type (e.g. diphtheria). For example, as specified in the cell marked as (a) in Row 7 of Figure 4, a child between 7-11 years old should receive his third dose at least *8* weeks after the second dose, and the vaccine should be *Td*, if the earlier dose constraint of Row 7 holds (his first diphtheria immunization was given before age of 1 year).

**Comment [MP2]:** R1.6

Figure 8 presents the complete specification of temporal constraints for the third diphtheria dose that take account of invalid doses. Readers who are not interested in the explanation of the complex immunization patterns shown in Figure 8 could move directly to the next section. Figure 8 extends Figure 4 in the following ways. First, its Row 5 contains additional information: it specifies the *interval flexibility* - the number of days of deviation from the minimum wait period or minimum age for administering the dose that is permitted for a dose to be counted as valid. In our case, if the third dose was given within 5 days before the minimum wait period or minimum age, then the dose is still considered valid.

**Comment [MP3]:** R1.1

Second, rows 9,11, 16, and 18 in the earlier-dose constraints part of Figure 8 shows that if any of the first 2 doses of Tdap that were given before the age of 7 years then they are invalid (due to *inadvertent use* of giving Tdap instead of the pediatric dose at an early age) and are not counted (NC). Therefore, in those rows, no information about the wait time and immunization type is provided in the right part of the table). Information about replacement doses is recorded in the tables for the doses that were missed (tables for doses 1 and 2). Row 22 in Figure 8 specifies the type and minimum waiting time for the replacement dose for dose 3 of Tdap given before the age of 7y (as specified in the left part of Row 22); as specified in the 3 last cells on Row 22, if after giving an invalid dose as a third

dose, you can immediately replace it, with zero wait time, with DTaP if the child is between 14 weeks and 7 years or with Td if the child is between 7 and 11 years or between 11 and 18 years old.

Comment [MP4]: R1.6

The *Current Age* and waiting period constraints appear in almost all immunization guidelines. The earlier-dose constraints are not always used. When they are used, the specification should completely cover all relevant combinations of dose number, type, and age interval. Note that more than one earlier-dose constraint may be applicable. In Figure 8, Row 10 shows that if the first dose was a non-Tdap that was given between the ages of 1 and 11y, and the child is currently over 11y, then the third dose to be given is not specified (this cell is marked with an 'a' in Figure 8). The third dose depends on the second dose given – whether it was Tdap or not, as specified by rows 15-21. Using the table, we can easily verify that rows 15 to 21 cover all combinations of dose type (Tdap or Non-Tdap) and the age range (between 1 and 18 years) for the second dose. Note that rows 15-16 refer to the second dose given before age of 1 year. Therefore, they do not complement Row 10, in which the first dose was given after age 1 year. This leaves the complementing cells to be the ones marked as b-f.

When more than one earlier-dose constraints are applicable and they specify different waiting periods and vaccines to give, we take the longest waiting period and permit either vaccine to be given. Rows 8 and 17 illustrate such a case. According to Row 8, if the first dose was a non-Tdap vaccine administered between 0 and 1 year of age, and the child is between 11 and 18 years old, then a Td dose can be given right away (see cell marked with 'h'). However, as specified by Row 17, if that child received non-Tdap vaccine as his *second* dose when he was between 1 and 11, then the third dose in the form of Tdap should be administered at least 6 months after the second dose (see cell marked with 'b'). Given these two constraints, we conclude that the third dose can be either Tdap or Td and should be given at least 6 months after the second dose.

Another example of multiple constraint permitting flexibility of vaccine administration is for the case where the first dose given was Tdap, it was administered between the ages of 7-11, and the child is now between 11-18y (Row 12). The immunization guideline tells us that because Tdap was a valid primary dose, then when the person receives an adult booster, the booster does not have to be Tdap.

To preserve this vaccine flexibility, Row 12 recommends Td as the third dose (see 'g' in Figure 8), but one of the rows 17, 19, 20, and 22, which concerns the second dose, are also applicable for this case:

- Rows 17 and 21 recommend Tdap for the third dose (if a non-Tdap was given as a second dose – see 'b' and 'f' in Figure 8)
- Rows 19 and 20 cover the case of the second dose also being Tdap (see 'd' and 'e' in Figure 8), and recommend Td as the third dose.

If certain physicians require a more strict interpretation of the guideline that does not offer the flexibility described above, then different versions of the immunization tables can be created, each representing a single strict interpretation that recommends only one type of vaccine per dose.

Additional constraints may refer to other interacting immunization types (e.g., MMR). Sometimes for the earlier-dose constraints that hold and for a certain age interval, the minimum wait time should be counted from the last dose of the vaccine type specified in the earlier-dose constraint. When more than one earlier-dose constraint applies, the earliest due date of vaccination is the date that satisfies all the relevant earlier dose constraints. When no previous dose constraints are specified, the wait time is counted from the last vaccine of the type being considered for vaccination.

We manually created instances in Protégé for immunization tables specified using the Immunization template, producing a knowledge base of immunization schedules. Figure 9 shows a screen shot of the diphtheria immunization schedule specified using the Protégé immunization ontology described in Section 2.2.4. As shown in Figure 2, the knowledge regarding immunization had to be collected not only from immunization tables found in the published guideline but also from textual recommendations and from footnotes; hence it would be very difficult to automate this step. However, the second step of translating our immunization table template into the Protégé format (such as that shown in Figure 9) could be automated in the future. This would save costs and prevent errors that might be introduced during a manual conversion process.

### 3.4. Validation of the immunization pattern's computability

To validate that the immunization template can be automatically-translated into code, we (1) wrote Jess rules for computing the due date of the next vaccination and (2) wrote an algorithm for translating an immunization pattern instance into a computer-interpretable guideline in the GLIF3 formalism.

3.4.1 Jess rules for computing the due date of the next vaccination

Jess (http://herzberg.ca.sandia.gov/jess/) is a rule engine and scripting environment written in Java. The Jess Tab in Protégé (http://www.ida.liu.se/~her/JessTab/) allows the integration of the Jess rule engine with a Protégé ontology. We wrote four Jess rules to compute the due date of the next vaccination. The rules operate on a simple virtual Medical Record (vMR) that has classes for holding a patient's demographics (name, ID, gender, date of birth), and previous substance administrations. The first rule, Determine-ImmunizationStatus, computes details of the most recent vaccinations of each vaccine type: vaccine, dose number, date, and age at the most recent vaccinations. The rule checks the vMR for invalid doses (i.e., inadvertent use and premature use). The criteria for inadvertent use are specified as an immunization constraint for a dose, such that if the constraint is satisfied, the dose should be repeated (normally, a constraint set for a dose would refer to previous doses rather than one that has the same dose number). Premature use is defined as deviating from the specified *interval flexibility* of the minimum wait period before doses and the minimum age for a given dose number. Doses that were found to be invalid are marked in the VMR as invalid and are not counted as valid doses. If the last dose was found to be inadvertent then the rule concludes an "inadvertent-use" assertion. The rule stores its results as instances of the class ImmunizationStatus.

Given the inferred ImmunizationStatus instances, the second rule, DetermineRelevantImmunizationConstraintSet, operates for each Immunization type. It first

consults the "inadvertent-use" status. If there was an inadvertent use, then the rule retrieves the relevant inadvertent use constraint for the next dose, searches for the immunization constraint for the next dose (dose i) in the ImmunizationConstraintSet for the next dose (dose i). Otherwise, the rule finds the age-appropriate earlier dose constraints. The rule stores its results as instances of the class AgeRelevantNextVaccinationConstraint.

The third rule, DetermineNextDueDate computes the next due date based on each AgeRelevantNextVaccinationConstraint, if they exists, or, if no age-appropriate earlier dose constraints exist, based just on current age and minimum-wait constraints. The rule stores its results as instances of the class RelevantNextVaccinationConstraint.

The fourth rule, DetermineTheLatest, computes a date for the next vaccination that satisfies all the RelevantNextVaccinationConstraint (maximum of the earliest date imposed by each earlier-dose constraint). The rule stores its results as instances of the class NextVaccination.

To evaluate the Jess rules, we created a simple vMR with several immunization records of four patients and verified that the Jess rules return due dates that correspond to the immunization templates that we instantiated in the knowledge base.

### 3.4.2 An algorithm for translating an immunization pattern into GLIF guidelines

The algorithm makes several assumptions when translating the immunization pattern into GLIF3 [46]:

(a) The decision criteria of GLIF decision steps use GEL [47] expressions without temporal operators so it would work with the GLIF execution engine – GLEE [48]; and (b) The GLIF encoding uses variable data items without worrying about how they are mapped to the actual EMR (e.g., Diphteria_date_of_last_dose, Diphteria_last_dose_num, 1_DTaP_DT_TDaP_date). Tools such as Knowledge-Data Ontology Mapper [49] can be used to write appropriate mapping functions.

The algorithm has three parts:

Part 1: creating the top-level GLIF guideline. The top-level algorithm consists of the following ordered Action Steps, as shown in Figure 10:

(a) Assigning possible vaccines – the step includes Assignment Action tasks for assigning lists of possible vaccine for each Immunization instance from the immunization template ontology discussed in Section 2.2.4 and exemplified in Figure 9.

(b) Get Data Action task for retrieving the patient's date of birth from the EMR

(c) Get Data Action tasks for retrieving the information about the last doses (dose number, vaccine, date, age at dose) from the EMR, accounting for invalid doses (these doses are not counted).

Then, a branch step is used to connect to different subguidelines – one for each Immunization instance from the immunization template ontology.

Part 2: Creating the subguideline for each immunization instance from the template ontology

The algorithm creates a decision step called "How many immunization doses given?". Based on the total number of doses specified in the immunization instance of the immunization template ontology, the appropriate number of decision options are created, each leading to an action step named "Potentially eligible for dose i", as shown in Figure 11. A patient is potentially eligible for dose i if he was given i-1 doses. The number of doses given is retrieved in Step 1(c) of the algorithm. The decision option that correspond to all doses having been administered leads to a patient state step named "done" (i.e., the patient is done).

Part 3: Creating the subguideline for each immunization dose from the template ontology

First a decision step "Current age?", is created. Then for each immunization dose constraint from the Immunization Template Ontology, we create decision options for the different age groups. Each decision option (marked as arcs in Figure 12) leads to a branch step from which a set of decision steps emanate. There is one decision step for each earlier-dose constraint. If the earlier-dose constraint is applicable then the earliest due date of the next dose can be computed based on the date of the last dose of the immunization specified by the earlier-dose dose constraint, the appropriating waiting time, and the recommended vaccine (step "Compute earliest due date"). If no earlier-dose dose constraint is

applicable, then the earliest due date is set to the date of the last dose (step "Earliest due date assigned with date of last dose"). Finally, for each age interval, the maximum of the due dates is taken to be the real due date for the next dose of the vaccine.

We checked the algorithm manually by creating examples of Diphteria instances of immunization doses 1, 2, 3 and MMR dose 2.

## 4. DISCUSSION

The so-called “knowledge-acquisition bottleneck”— the difficulty of transforming knowledge from the forms in which it is available into forms that can be used by a knowledge-based system—plagued developers of clinical DSS from the early days [50]. Guideline formalisms developed in the last two decades tried to create computer-interpretable models of clinical guideline so that the abstractions and recommendations in a guideline would not have to be encoded in generic rule or procedural languages. Nevertheless, these formalisms still aim for generality in being able to encode all kinds of guidelines. Accordingly, the primitives provided in these formalisms are abstractions such as actions, decisions, and criteria. Clinicians and knowledge engineers still have to conceptualize a guideline in these terms [51]. Work in the 1980's had shown the power of domain-specific models. For example, OPAL, a domain-specific tool for acquiring cancer chemotherapy protocols, was shown to be usable by clinicians [52]. More recently, in software engineering, advocates of *domain-specific modeling* have argued for the productivity to be gained from using high-level domain-specific abstractions [53]. Our work on guideline design patterns aims to provide domain experts with abstractions that are defined at the same level as that of their guidelines. To check the robustness of our patterns, we used them to specify a substantial number of guidelines.

Another advantage of our templates for screening and immunization guideline is the possibility to check completeness and consistency of encoded guidelines. Miller and co-authors [23] noted that for screening the validity of prior doses, they need a small amount of logic that could be represented in a decision table. But for computing the next due date, as we are doing using generic Jess rules that operate on declarative knowledge structures, they found that they had to represent relevant clinical knowledge as specific if-then-else rules because of the large number of clinical parameters and their

combinations. Their goal was to express in tables as much of the logic that is frequently updated as possible. However, much of the logic still resides in if-then-else rules. In contrast, our representation enables us to represent different types of temporal constraints that occur in immunization guidelines in a single table. The tabular representation facilitates the specification of complete and consistent immunization schedules even without using elaborate methods for temporal constraints checking, such as [54]. However, it must be acknowledged that we deliberately chose to focus on the specification of temporal constraints, leaving other conditions, such as contraindications, risks, gender, occupation, and place of residence to be represented by some other expression language.

The specificity of domain-specific models has disadvantages as well. Their lack of generality means that our templates are suitable only for screening and immunization guidelines. Thus, until patterns are defined for additional domains, guidelines for which patterns have not been defined must be represented using generic guideline-modeling primitives.

There are evidences that support the existence of patterns in other areas of medicine. Clinical trials, for example, have very stereotypical structures that developers of clinical trial ontologies have taken advantage of [55].BRIDG, a domain analysis model adopted by NCI, HL7 and CDISC, for example, specifies in great detail the components of clinical trials [56]. TrialWiz, a tool for authoring clinical trials for the Immune Tolerance Network, uses an ontology of clinical trials to manage the complexity of the protocol-encoding process [57]. Likewise, radiology guidelines follow patterns that describe the appropriateness criteria for deciding on different radiology modalities, based on clinical conditions and their variants and the ratings of possible radiology modalities [58]. Within the guideline modeling domain, the PRODIGY3 system focused their models on chronic disease management by primary care physicians [34]. The PRODIGY3 formalism can be seen as a complex template for such guidelines. It structures a chronic-disease management guidelines as a structured collection of decisions, to be made by general practitioners, organized in Decision Maps –collections of clinical contexts with associated management decisions. The clinical contexts are defined as logical conditions that relate to patient risks, comorbid conditions that affect treatment, current medications taken, and whether guideline goals are met (e.g., possible clinical contexts for a high blood cholesterol

guideline include "no drug, high Congestive Heart Disease risk", "Drug therapy, 2+ drugs"). For each context, a different map path is defined. Once a decision about a drug regime (e.g., ACE Inhibitor) is made, PRODIGY-3 provides templates for refining the drug regime to a generic drug (e.g., lisinopril), and then to specific prescriptions.

The patterns we have defined for screening and immunization guidelines can be complemented with patterns that other researchers have found. For example, we could integrate MHB's structural patterns [17] for sections of guidelines to help us create definitions, qualitative facts, and relations between terms. The SAGE [19] expression templates would be useful for encoding indications and contraindications of vaccines. Similarly, our design patterns could complement other methodologies for formalizing guidelines. For example, within the methodology advocated by Shiffman and colleagues [59], these design patterns can provide the overarching framework within which the semantic refinement (atomization, de-abstraction, and disambiguation) take place. Furthermore, the immunization templates easily facilitate the verification of completeness and consistency as specified in Shiffman's methodology.

Future experimentation should evaluate whether the design patterns proposed actually help inexperienced modelers in creating correct guideline models more quickly (overcoming "analysis paralysis") and generate models that are conforming to the patterns. Creating standard models is important for novices; standard solutions that adhere to some quality-control standard, such as the check list that we developed for the screening guidelines, are preferred to custom-made models that may sometimes be better than the standard solution, but other times may be worse.

It would also be interesting to assess whether the results could be attributed just to the conceptualization used in the templates rather than being dependent on the syntax of the templates or the tool in which the templates are constructed (Protégé). We could test whether teaching students the conceptualization of screening templates and then asking them to model screening guidelines directly in a guideline modeling language such as GLIF3 would achieve similar effects to those achieved using the templates directly for modeling. Results from an experiment in analogical problem solving

conducted by Gick and Holyoak [60] demonstrated that people can develop a solution to a problem by using an analogous problem that they recall from memory and its solution, provided they are given a hint to use the analogous problem to help solve the problem. If indeed, the templates could be used as an analogy to guide direct representation in a guideline modeling language such as GLIF3, then there would be no need to create mapping tools that would convert encoded templates into a guideline modeling language. In this way, a modeler could benefit from the conceptualization made easy by the template and at the same time could execute the resulting encoding in a guideline execution engine that is available for that guideline modeling language. Since the patterns that we have developed are independent of any guideline modeling language, they could potentially be used to assist conceptualization in any guideline modeling language.

## 5. CONCLUSION

Design patterns for clinical guidelines are promising tools for facilitating the formalization process of narrative guidelines in specific domains. The screening and immunization templates could potentially be applied to encode any guidelines belonging to these categories, as demonstrated by our ability to encode all of the screening guidelines from the National Guideline Clearinghouse that contained a clinical algorithm and all of the CDC immunization guidelines. In addition, knowledge modeled in the immunization patterns (1) can be checked for completeness and consistency, (2) could be used by applications that execute workflow processes or that compute the due date of the next vaccination, and (3) could be automatically translated into a computer-interpretable guideline.

**Acknowledgments**

We would like to thank Prof. Perry Miller, Dr. Sandra Frawley for their detailed comments, ideas, and clarifications that helped us to understand issues related to immunization considerations and improve

our formalization of immunization constraints, and Daniel Rubin for his helpful suggestions. We thank the anonymous reviewers for their helpful suggestions.

**Comment [MP5]:** Revised per editor's comments

## REFERENCES


1. M.J. Field, K.N. Lohr. *Guidelines for Clinical Practice: Directions for a New Program* (Washington DC: Institute of Medicine, National Academy Press; 1990).
2. S. Shea, W. DuMouchel, L. Bahamonde. A Meta-analysis of 16 Randomized Controlled Trials to Evaluate Computer-based Clinical Reminder Systems for Preventative Care in the Ambulatory Setting. *J Am Med Inform Assoc* 3 (1996) 399-409.
3. M. Peleg. Guideline and Workflow Models, in: Greenes RA, ed., *Clinical Decision Support - The Road Ahead* (Elsevier/Academic Press, Amsterdam; Boston, 2006).
4. P.A. de-Clercq, J.A. Blom, H.H. Korsten, A. Hasman. Approaches for creating computer-interpretable guidelines that facilitate decision support. *Artif Intell Med* 31(2004) 1-27.
5. W.M. Tierney, J.M. Overhage, B.Y. Takesue, L.E. Harris, M.D. Murray, D.L. Vargo, et al. Computerizing guidelines to improve care and patient outcomes: the example of heart failure. *J Am Med Inform Assoc* 2(1995) 316-22.
6. E. Shalom, Y. Shahar, M. Taieb-Maimon, G. Bar, A. Yarkoni, O. Young, et al. A quantitative assessment of a methodology for collaborative specification and evaluation of clinical guidelines. *J Biomed Inform* 41(2008) 889-903.
7. R.N. Shiffman, J. Dixon, C. Brandt, A. Essaihi, A. Hsiao, G. Michel, et al. The GuideLine Implementability Appraisal (GLIA): development of an instrument to identify obstacles to guideline implementation. *BMC Med Inform Decis Mak* 5 (2005) 23.
8. R.N. Shiffman, G. Michel, A. Essaihi, T.W. Marcy. Using a guideline-centered approach for the design of a clinical decision support system to promote smoking cessation. *Stud Health Technol Inform* 101 (2004)152-6.
9. Y. Shahar, O. Young, E. Shalom, M. Galperin, A. Mayaffit, R. Moskovitch, et al. A framework for a distributed, hybrid, multiple-ontology clinical guideline library, and automated guideline-support tools. *J Biomed Inform* 37 (2004) 325-44.
10. P. Votruba, S. Miksch, R. Kosara. Facilitating Knowledge Maintenance of Clinical Guidelines and Protocols. *Stud Health Technol Inform* 107 (2004) 57-61.
11. A. ten-Teije, M. Marcos, M. Balser, J.v. Croonenborg, C. Duelli, F.v. Harmelen, et al. Improving medical protocols by formal methods. *Artif Intell Med* 36 (2006) 193-209.
12. M. Peleg, S. Tu, A. Mahindroo, R. Altman. Modeling and Analyzing Biomedical Processes using

Workflow/Petri Net Models and Tools, *Stud Health Technol Inform* 107 (2004) 74-8.

13. A.A. Boxwala, P. Mehta, M. Peleg, R. Lacson, N. Ash, J. Bury, et al. Modeling guidelines using domain-level knowledge representation components, in: J.M. Overhage, editor, P*roc AMIA Symp*. (American Medical Informatics Association, Los Angeles, 2000) 974.
14. R. Serban, A. ten Teije, F. van Harmelen, M. Marcos, C. Polo-Conde. Extraction and use of linguistic patterns for modelling medical guidelines, *Artif Intell Med* 39 (2007) 137-49.
15. A. Essaihi, G. Michel, R.N. Shiffman. Comprehensive categorization of guideline recommendations: creating an action palette for implementers, in: M.A. Musen, editor *Proc AMIA Symp* (American Medical Informatics Association, Washington D.C., 2003) 220-4.
16. E. Gamma, R. Helm, R. Johnson, J.M. Vlissides. *Design patterns: Elements of reusable object-oriented software*, (Reading, MA, Addison-Wesley Publishing Company, 1995).
17. A. Seyfang, S. Miksch, M. Marcos, J. Wittenberg, C. Polo-Conde, K. Rosenbrand. Bridging the Gap between Informal and Formal Guideline Representations, in: G. Brewka, S. Coradeschi, A. Perini, P. Traverso, editors, *European Conference on Artificial Intelligence*, (IOS Press, Riva del Garda, 2006).
18. Y. Shahar, S. Miksch, P. Johnson. The Asgaard Project: A Task-Specific Framework for the Application and Critiquing of Time-Oriented Clinical Guidelines. *Artificial Intelligence in Medicine* 14 (1998) 29-51.
19. S.W. Tu, J.R. Campbell, J. Glasgow, M.A. Nyman, R. McClure, P.C. J McClay J, Hrabak KM, Berg D, Weida T, Mansfield JG, Musen MA, Abarbanel RM. The SAGE Guideline Model: achievements and overview. *J Am Med Inform Assoc* 14 (2007) 589-98.
20. M. Sordo, O. Ogunyemi, A.A. Boxwala, R.A. Greenes, S. Tu. Software Specifications for GELLO: An Object-Oriented Query and Expression Language for Clinical Decision Support: Decision Systems Group Report DSG-TR-2003-02, (2004).
21. M. Peleg, A. Boxwala, O. Ogunyemi, Q. Zeng, S. Tu, R. Lacson, et al. GLIF3 Technical Documentation, in: http://smi.stanford.edu/projects/intermed-web/ (Accessed: 14 April, 2009)
22. P.L. Miller, S.J. Frawley, F.G. Sayward. Maintaining and Incrementally Revalidating a Computer-Based Clinical Guideline: A Case Study. *J Biomed Inform* 34 (2001) 99-111.
23. P.L. Miller, S.J. Frawley, F.G. Sayward, W.A. Yasnoff, L. Duncanc, D.W. Fleming. Combining Tabular, Rule-Based, and Procedural Knowledge in Computer-Based Guidelines for Childhood Immunization. *Computers and Biomedical Research* 30 (1997) 211-231.
24. D. Wang, R.A. Jenders, B. Dasgupta. Translating national childhood immunization guidelines to a

computer-based reminder recall system within an immunization registry, in: N.M. Lorenzi, editor, *Proc AMIA Symp* (American Medical Informatics Association, Washington D.C., 1999) 994-8.

25. V. Anand, P.G. Biondich, G. Liu, M. Rosenman, S.M. Downs. Child Health Improvement through Computer Automation: the CHICA system, Stud Health Technol Inform (2004) 187-91.
26. J. Dulcey. *Micromedex medical logic modules* (The Thomson Cooperation, 2002) http://www.ids-healthcare.com/Common/Paper/Paper_185/mlm_whitepaper.pdf (Accessed: 14 April, 2009).
27. J. Choi, S. Bakken. Comparison of Primary Care Expert and Computer-Interpretable Depression Screening Guideline Recommendations, in: D. W. Bates, editor, *Proc. AMIA Symp* (American Medical Informatics Association, Washington D.C., 2006) 887.
28. J. Fox, V. Patkar, R. Thomson, E. Black. PROforma guidelines and care pathways: performance in trials and future plans, in: Annette ten Teije , P. Lucas, S. Miksch, editors, *The biennial European Conference on Artificial Intelligence (ECAI) Workshop: AI techniques in healthcare: computerized guidelines and protocols* (IOS Press, Riva del Garda, Italy, 2006).
29. F. Chesani, P.D. Matteis, P. Mello, M. Montali, S. Storari. A Framework for Defining and Verifying Clinical Guidelines: A Case Study on Cancer Screening, in: F. Esposito, Z. W. Ras, D. Malerba, G. Semeraro, Editors , *Lecture Notes in Computer Science* 4203 (Springer Berlin / Heidelberg 2006) 338-343.
30. N.-J. Lee, R. John, S. Bakken. Functional Requirements Specification and Data Modeling for a PDA-based Decision Support System for the Screening and Management of Obesity. in: D.W. Bates, editor, *Proc. AMIA Symp* (American Medical Informatics Association, Washington D.C., 2006) 1002.
31. M. Peleg, A.A. Boxwala, E. Bernstam, S. Tu, R.A. Greenes, E.H. Shortliffe. Sharable Representation of Clinical Guidelines in GLIF: Relationship to the Arden Syntax. *Journal of Biomedical Informatics* 34 (2001) 170-81.
32. J. Choi, A.A. Boxwala. Representing domain-level knowledge components using primitives, in: Suzanne Bakken, editor, *Proc AMIA Ann Symp* (American Medical Informatics Association, Washington D.C., 2001) 879.
33. S.W. Tu, J.R. Campbell, M.A. Musen. SAGE Guideline Modeling: Motivations and Methodology, in: K. Kaiser, S. Miksch, and S. W. Tu,, editors, *Proceedings of the Symposium on Computerized Guidelines and Protocols* (IOS Press, Amsterdam,, 2004) 167-171.
34. P.D. Johnson, S.W. Tu, N. Booth, B. Sugden, I.N. Purves. Using Scenarios in Chronic Disease Management Guidelines for Primary Care, in: J.M. Overhage, editor, P*roc AMIA Symp*. (American Medical Informatics

Association, Los Angeles, 2000) 389-393.

35. M. Peleg, S.W. Tu, J. Bury, P. Ciccarese, J. Fox, R.A. Greenes, et al. Comparing Computer-Interpretable Guideline Models: A Case-Study Approach., *J Am Med Inform Assoc* 10 (2003) 52-68.
36. E. Yourdon. Chapter 19. Building a Preliminary Behavioral Model, in: *Just Enough Structured Analysis*, http://yourdon.com/strucanalysis/wiki, (Accessed: 14 April, 2009)
37. W.E. Grosso, H. Eriksson, R. Fergerson, J.H. Gennari, S.W. Tu, M.A. Musen. Knowledge Modeling at the Millennium (The Design and Evolution of Protege-2000), in: Gains BR, Kremer R, Musen M, eds. *The 12th Banff Knowledge Acquisition for Knowledge-Based Systems Workshop* (Banff, Canada; 1999) 7-4-1 to 7-4-36.
38. T.R. Gruber. Toward Principles for the Design of Ontologies Used for Knowledge Sharing. *Intl. Journal of Human-Computer Studies* 43 (1995) 907-928.
39. American College of Foot and Ankle Surgeons. Diabetic foot disorders: a clinical practice guideline, *J Foot Ankle Surg* 39(2000) S1-60.
40. US Dept. of Veterans Affairs. VA/DODclinical practice guideline for the management of diabetes mellitus in primary care. *Module E - eye care* (US Dept. of Veterans Affairs, Washington D.C., 2003).
41. R.S. Braithwaite, J. Concato, C.C. Chang, M.S. Roberts, A.C. Justice. A Framework for Tailoring Clinical Guidelines to Comorbidity at the Point of Care, *Arch Intern Med* 167 (2007) 2361-2365.
42. STARD Group. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Clin Chem* 49 (2003) 1-6.
43. Centers for Disease Control and Prevention. General recommendations on immunization recommendations of the Advisory Committee on Immunization Practices (ACIP). MMWR Morb Mortal Wkly Report, 43(1994) 1-38.
44. Institute for Clinical Systems Improvement (ICSI). Health Care Guideline: Immunizations, Twelfth Edition, in: http://www.icsi.org/immunizations___guideline_/immunizations__guideline__38400.html, (ICSI, Accessed 14 April, 2009)
45. Y. Shahar, D. Goren-Bar, D. Boaz, G. Tahan. Distributed, intelligent, interactive visualization and exploration of time-oriented clinical data and their abstractions, *Artif Intell Med* 38 (2006) 115-35.
46. A.A. Boxwala, M. Peleg, S. Tu, O. Ogunyemi, Q. Zeng, D. Wang, et al. GLIF3: a representation format for sharable computer-interpretable clinical practice guidelines. *Journal of Biomedical Informatics* 37 (2004) 147-61.

47. M. Peleg, O. Ogunyemi, S. Tu, A.A. Boxwala, Q. Zeng, R.A. Greenes, et al. Using Features of Arden Syntax with Object-Oriented Medical Data Models for Guideline Modeling, in: Suzanne Bakken, editor, *Proc AMIA Ann Symp* (American Medical Informatics Association, Washington D.C., 2001) 523-7.
48. D. Wang, M. Peleg, S.W. Tu, A.A. Boxwala, O. Ogunyemi, Q. Zeng, et al. Design and implementation of the GLIF3 guideline execution engine. *Journal of Biomedical Informatics* 37 (2004) 305-18.
49. M. Peleg, S. Keren, Y. Denekamp. Mapping Computerized Clinical Guidelines to Electronic Medical Records: Knowledge-Data Ontological Mapper (KDOM). *Journal of Biomedical Informatics* 41 (2008) 180-201.
50. M.A. Musen. Widening the knowledge-acquisition bottleneck: Automated tools for building and extending clinical methods, in: W.E. Hammond, editor, Proceedings of the *Congress* on Medical Informatics *(AAMSI)* (American Medical Informatics Association, San Francisco, 1989) 2-7.
51. E. Shalom, Y. Shahar, E. Lunenfeld, M. Taieb-Maimon, O. Young, D. Goren-Bar, et al. The Importance of Creating an Ontology-Specific Consensus Before a Markup-Based Specification of Clinical Guidelines, in: Annette ten Teije , P. Lucas, S. Miksch, editors, *Proceeding of the biennial European Conference on Artificial Intelligence* (IOS Press, Riva del Garda, Italy, 2006).
52. M.A. Musen, L.M. Fagan, D.M. Combs, E.H. Shortliffe. Use of domain model to drive an interactive knowledge-editing tool. *International Journal of Man-Machine Studies* 26 (1987) 105-121.
53. S. Kelly, J.P. Tolvanen. *Domain-specific modeling: Enabling full code generation*, (Wiley-IEEE Computer Society Press, Alamitos, California, 2008).
54. G. Duftschmid, S. Miksch, W. Gall. Verification of temporal scheduling constraints in clinical practice guidelines, *Artif Intell Med* 25 (2002) 93-121.
55. R. Shankar, M. O'Connor, D.B. Parrish, S.B. Martins, A.K. Das. Epoch: An ontological framework to support clinical trials management, in: L. Xiong, Y. Xia, editors, *International Workshop on Health Information and Knowledge Management* (ACM, Ney York, 2006) 25-32.
56. D.B. Fridsma, J. Evans, S. Hastak, C.N. Mead. The BRIDG Project: A Technical Report. *J Biomed Inform* 15 (2008) 130–137.
57. R. Shankar, S. Arkalgud, M. O'Connor, K. Boyce, D. Parrish, A. Das. TrialWiz: an Ontology-Driven Tool for Authoring Clinical Trial Protocols, in: J. Suermondt, *Proc AMIA Symp* (American Medical Informatics Association, , Washington D.C. 2008) 1226.
58. D. Rubin. Templates for appropriateness criteria in radiology, personal communication, (Stanford, CA,

2008)

59. R.N. Shiffman, G. Michel, A. Essaihi, E. Thornquist. Bridging the guideline implementation gap: a systematic, document-centered approach to guideline implementation, *J Am Med Inform Assoc* 11 (2004) 418-26.

60. M.L. Gick, K.J. Holyoak. Analogical Problem Solving, *Cognitive Psychology* 12 (1980) 306-355.

**Figure and table Captions**

**Figure 1.** A check list for improving modeling of screening guidelines using screening patterns

**Figure 2.** A portion of the childhood immunization guideline [44] of the Institute for Clinical Systems Improvement (ICSI). The box in the catch-up immunization schedule table specifies the text that is copied to the box on its right. These are directions for the minimum interval between doses 1 and 2 of Pneumococcal vaccine

**Figure 3.** SAGE representation for giving the third dose of PCV7 (Pneumococcal vaccine). The criteria that must all hold in order to give the third PCV7 dose are shown in the frame on the left. The last criterion is expanded in the frame on the top right. They all relate to the time that has passed since the last dose given (the second dose). The highlighted criterion in the top right frame is expanded in the bottom frame. The bottom frame shows the use of a generic temporal-comparison template to encode the statement “the most recent Pneumococcal 7 valent conjugate Substance Administration was done more than 28 days ago (before now)”

**Figure 4**. A simplified immunization table for the third dose of diphtheria immunization, in Excel.

**Figure 5**. A UML class diagram showing the main classes in the Immunization Template Ontology

**Figure 6.** A screen shot of the cervical cancer screening guideline using the screening template in Protégé. The diagram shows that when the time for screening for cervical cancer has arrived, data for screening is collected and patients are categorized into two possible categories: "Cancer Management" if [physical_crevix_exam = positive or PAP = positive] or "Continue Screening" if [physical_crevix_exam = negative and PAP = negative].

**Figure 7.** A screen shot of the colon cancer data collection template in Bonita Workflow Editor. Selecting Test represents a choice between three tests that can be performed: FOBT (if condition C1 holds), Sigmoidoscopy (if condition C2 holds) and Colonoscopy (if condition C3 holds). Following each test is a step that inspects test results once the results are available. In the case of inspecting results of FOBT (Inspecting_Results1), the results may be positive (C4), in which case, Colonoscopy

is done or negative, in which case the workflow ends. In the case of inspecting Sigmoidoscopy results (Inspecting_Results2) the results may be inadequate (C5a) in which case Sigmoidoscopy is repeated, inconclusive (C5b) in which case colonoscopy is done or conclusive, in which case the workflow ends. In the case of inspecting colonoscopy results (Inspecting_Results3), the results can be inadequate due to inadequate preparation (C6a) in which case the colonoscopy (which includes cleaning) is repeated, inadequate but the cleaning was adequate (C6b), in which case Colonography is done and upon arrival of results the end of the workflow is reached (this is according to the guideline, which did not specify other cases that depend on colonography results), or adequate, in which case the workflow ends. The following preconditions are used: c1 [FOBT]; c2 [sigmoidoscopy]; c3 [colonoscopy]; c4 [FOBT_result = positive]; c5a [sigmoidoscopy_result = inadequate]; c5b [sigmoidoscopy_result = inconclusive]; c6a [colonoscopy_result inadequate and preparation inadequate]; c6b [colonoscopy_result inadequate and preparation adequate]; c6c [colonoscopy_result adequate].

**Figure 8.** The complete immunization table for the third dose of diphtheria immunization, in Excel. NA – not applicable; NC – not counted.

**Figure 9**. A screen shot of the diphtheria immunization schedule specified using Protégé. An instance of the Immunization class is shown on top. The constraints for the third dose are shown on the bottom as an instance of the *Immunization Dose Constraints* Class, which contains the four other Immunization Constraints: Immunization Constraint Set, Current Age Constraint, Next Vaccination Constraint, and Earlier Dose Constraint. For example, the first dose constraint states that if there were 2 previous doses given, and the current age is between 14 weeks and 7 years, then if the first (earlier) dose of either DTaP, or Td was given when the patient was between 0 and 1 year old, then the recommended vaccine is DTaP and should be given at least 4 weeks after the previous dose.

**Figure 10.** An example of the top-level GLIF algorithm generated from the immunization pattern. Action steps are depicted as green rectangles; branch step as an orange oval.

**Figure 11.** The top-level subguideline for the Diphtheria immunization generated from the immunization pattern. Blue hexagon depicts a decision step; yellow diamond – patient state step.

**Figure 12.** Part of a GLIF algorithm for dose 3 of the diphtheria immunization. The part shown is only for the age group 14weeks – 7 years.

**Table 1. Characteristics of various approaches for design patterns for clinical guidelines**

**Table 2.** time-for-screening table for the cervical cancer screening guideline

**Table 1. Characteristics of various approaches for design patterns for clinical guidelines**

| Design pattern<br>Characteristic | GLIF Macro [13, 21] | MHB [17] | Linguistic Pattern [14] | Action Palette [15] | Expression Templates [19] |
|---|---|---|---|---|---|
| **Purpose** | Ease conceptualization | Ease migration from narrative to encoding | Ease migration from narrative to encoding | Standardization | Ease formalization |
| **Scope** | Recommendation set | Individual recommendation | Individual recommendation | Part of individual recommendation | Part of individual recommendation |
| **Clinical vs. data-based semantics** | Clinical | Clinical and data-based | Clinical | Clinical | Data-based |
| **Specifies entire process of care** | Yes | No | No | No | No |
| **Specifies control-flow** | Yes | Yes | No | No | No |
| **Visual pattern** | Yes | No | No | No | No |
| **Automated vs. user-assisted representation** | User | User | automatic | User | User |

**Table 2.** Time-for-screening table for the cervical cancer screening guideline

| data_item | frequency | criterion |
|---|---|---|
| PAP_smear | 6-12m | physical_cervix_exam = negative and PAP = negative and HPV_DNA = positive |
| PAP_smear | 1-2y | physical_cervix_exam = negative and PAP = negative and HPV_DNA = unknown and age >=30 |
| PAP_smear | 2y | physical_cervix_exam = negative and PAP = negative and ( age <30 or (age >=30 and HPV_DNA = negative) |

**Screening Pattern Check List**

- Risk categories are modeled explicitly
- The number of risk categories is minimal and categories do not overlap
- Conditions for risk categories are defined
- Conditions refer to defined variables (variables that are read from the EMR, or which were assigned with expressions relating to EMR values. Do not write statements such as "*result* = ill" without defining the variable result)
- Screening frequency is updated after risk category is established
- Patient education step exists[1]
- Screening pattern does not elaborate data collection

[1] Some clinical guidelines do not include a patient education step

**Data Collection Check List**

- All alternative test results are covered (including those that are implicit in the clinical guideline. e.g., result is neg/pos, result inadequate/adequate)
- Conditions for test selection are provided
- Conditions for test selection refer to:
  - The test selected
  - Screening interval
- Conditions refer to defined variables (variables that are read from the EMR, or which were assigned with expressions relating to EMR values. Do not write statements such as "*result* = ill" without defining the variable result)
- No loose ends left (all paths converge at a single end point)
- The Data Collection pattern does not establish risk categories

**Figure 1**

**Normal Schedule**

| Vaccine | Birth | 1 mo | 2 mo | 4 mo | 6 mo | 12 mo | 15 mo | 18 mo | 24 mo | 4-6 yr | 11-12 yr | 15-18 yr |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. DTaP | | | X | X | X | X | | | | X | Tdap | |
| 2. IPV | | | X | X | X | | | | | X | | |
| 3. MMR (MMRV) | Combined measles, mumps, rubella and varicella vaccine (MMRV) is preferred for children 12 months through 12 years of age over separate injection of equivalent component vaccines. | | | | | X | | | | X | | |
| 4. Varicella | | | | | | X | | | | X | | X, verify second dose completed |
| 5. Pneumococcal (PCV7) | | | X | X | X | X | | | | | | |

**Catch-up Immunization Schedule**
**for Persons Aged 4 Months–18 Years Who Start Late or Who Are More Than 1 Month Behind**

UNITED STATES • 2007

The table below provides catch-up schedules and minimum intervals between doses for children whose vaccinations have been delayed. A vaccine series does not need to be restarted, regardless of the time that has elapsed between doses. Use the section appropriate for the child's age.

**CATCH-UP SCHEDULE FOR PERSONS AGED 4 MONTHS–6 YEARS**

| Vaccine | Minimum Age for Dose 1 | Minimum Interval Between Doses: Dose 1 to Dose 2 | Dose 2 to Dose 3 | Dose 3 to Dose 4 | Dose 4 to Dose 5 |
|---|---|---|---|---|---|
| Diphtheria, Tetanus, Pertussis[3] | 6 wks | 4 weeks | 4 weeks | 6 months | 6 months[3] |
| *Haemophilus influenzae* type b[4] | 6 wks | 4 weeks<br>If first dose administered at age <12 months<br>8 weeks (as final dose)<br>If first dose administered at age 12-14 months<br>No further doses needed<br>If first dose administered at age ≥15 months | 4 weeks[4]<br>If current age <12 months<br>8 weeks (as final dose)[4]<br>If current age ≥12 months and second dose administered at age <15 months<br>No f<br>If previous | 8 weeks (as final dose)<br>This dose only necessary for children aged 12 months–5 years who received 3 doses | |
| Pneumococcal[5] | 6 wks | 4 weeks<br>If first dose administered at age <12 months and current age <24 months<br>8 weeks (as final dose)<br>If first dose administered at age ≥12 months or current age 24–59 months<br>No further doses needed<br>for healthy children if first dose administered at age ≥24 months | 8 v<br>No f<br>for he<br>adm | | |
| **CATCH-UP SCHEDULE FOR PERSON** | | | | | |
| Tetanus, Diphtheria/ Tetanus, Diphtheria, Pertussis[10] | 7 yrs[10] | 4 weeks | If first dos<br>If first dose administered at age ≥12 months | age <12 months | |

4 weeks
If first dose administrated at age < 12 months and current age < 24 months

8 weeks (as final dose)
If first dose administrated at age > 12 months or current age 24-59 months

No further doses needed
For healthy children if first dose administered at age ≥ 24 months

**Footnotes**

**Tetanus and diphtheria toxoids vaccine (Td) and tetanus and diphtheria toxoids and acellular pertussis vaccine (Tdap).**
(Minimum ages: 7 years for Td, 10 years for BOOSTRIX®, and 11 years for ADACE™)
Tdap should be substituted for a single dose of Td in the primary catch-up series or as a booster if age appropriate; use Td for other doses.

**Figure 2**

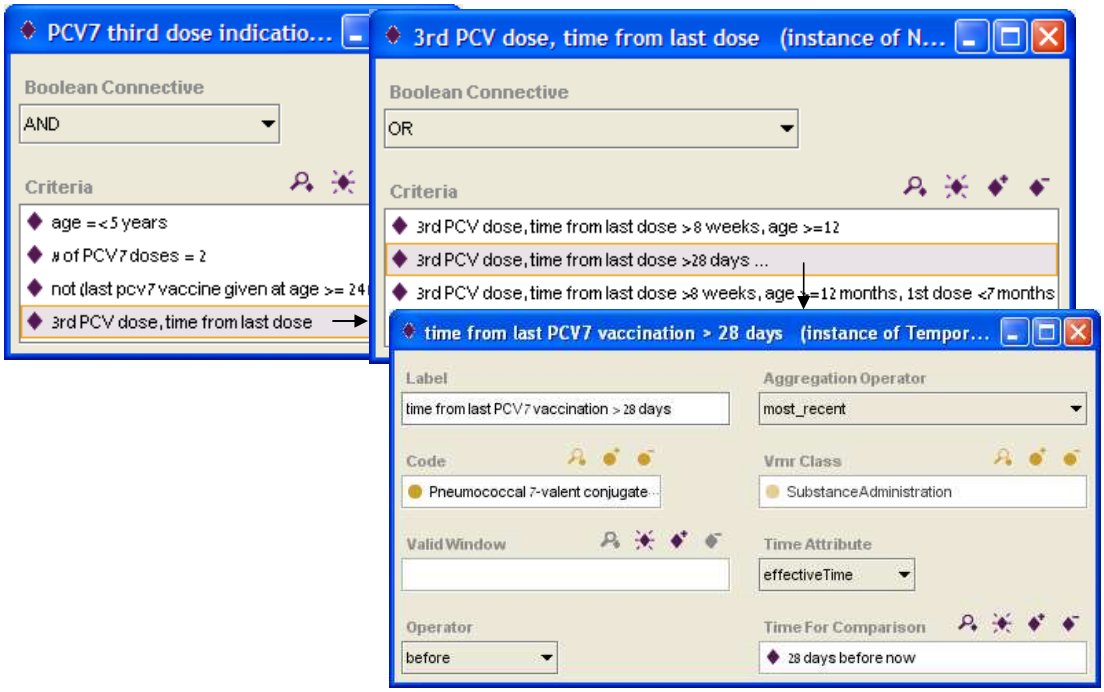


**Figure 3**

| Row | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 1 | Vaccines | | DTaP, Td, Tdap | | | | | |
| 2 | *Relevant prior vaccines | | DTaP, Td, Tdap | | | | | |
| 3 | Dose number | | 3 | | | | | |
| 4 | Recommended age | | 6m | | | | | |
| 5 | Earlier dose constraints: | | age when dose was given | | Current age | | | |
| 6 | Type | Dose# | Min age | Max age | (0,14w] | (14w, 7y] | (7y, 11y] | (11y-18y] |
| 7 | * | 1 | 0 | 1y | - | 4w DTaP | 8w Td ⓐ | 8w Tdap |
| 8 | * | 1 | 1y | 18y | - | 4w DTaP | 6m Td | 6m Tdap |

Any affecting vaccine (DTaP, Td, Tdap)

Vaccine
Minimum waiting time since previous dose

**Figure 4**

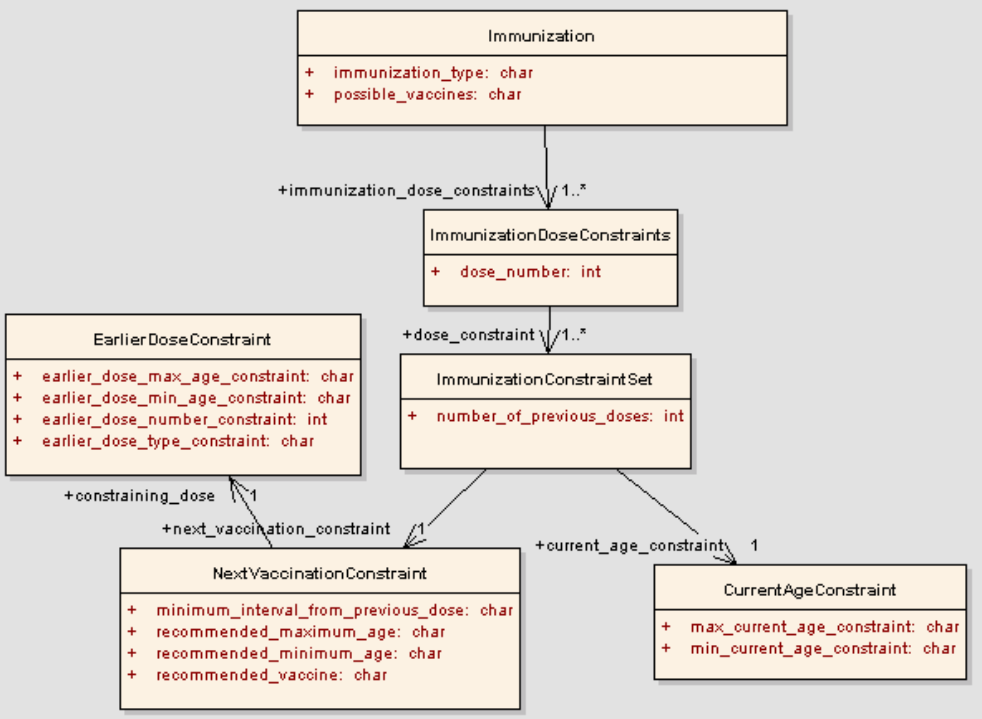
Immunization
+ immunization_type: char
+ possible_vaccines: char
+immunization_dose_constraints
1..*
ImmunizationDoseConstraints
+ dose_number: int
+dose_constraint
1..*
EarlierDoseConstraint
+ earlier_dose_max_age_constraint: char
+ earlier_dose_min_age_constraint: char
+ earlier_dose_number_constraint: int
+ earlier_dose_type_constraint: char
ImmunizationConstraintSet
+ number_of_previous_doses: int
+constraining_dose
1
+next_vaccination_constraint
1
+current_age_constraint
1
NextVaccinationConstraint
+ minimum_interval_from_previous_dose: char
+ recommended_maximum_age: char
+ recommended_minimum_age: char
+ recommended_vaccine: char
CurrentAgeConstraint
+ max_current_age_constraint: char
+ min_current_age_constraint: char

**Figure 5**

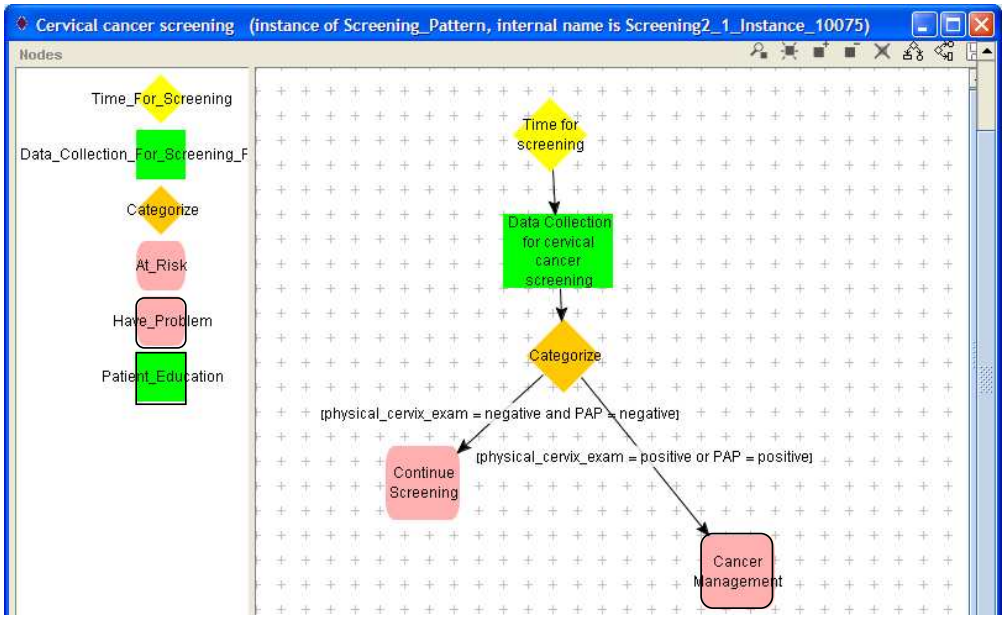


**Figure 6**

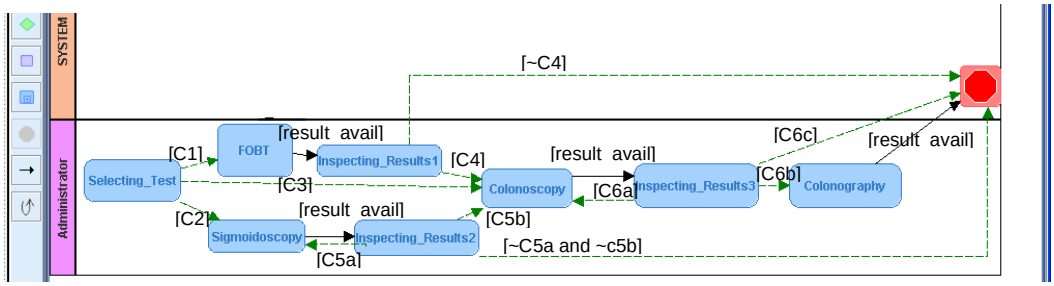


**Figure 7**

| Row | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 1 | Vaccines | | DTaP, Td, Tdap | | | | | |
| 2 | *Relevant prior vaccines | | DTaP, Td, Tdap | | | | | |
| 3 | Dose number | | 3 | | | | | |
| 4 | Recommended age | | 6m | | | | | |
| 5 | Interval flexibility | | 5d | | | | | |
| 6 | Earlier dose constraints: | | age when dose was given | | Current age | | | |
| 7 | Type | Dose# | Min age | Max age | (0,14w] | (14w, 7y] | (7y, 11y] | (11y-18y] |
| 8 | Non-Tdap | 1 | 0 | 1y | - | 4w DTaP | 8w Td | 0 Td [h] |
| 9 | Tdap | 1 | 0 | 1y | NC | NC | NC | NC |
| 10 | Non-Tdap | 1 | 1y | 11y | NA | 4w DTaP | 6m Td | - [a] |
| 11 | Tdap | 1 | 1y | 7y | NC | NC | NC | NC |
| 12 | Tdap | 1 | 7y | 11y | NA | NA | 6m Td | 6m Td [g] |
| 13 | Tdap | 1 | 11y | 18y | NA | NA | NA | 6m Td |
| 14 | Non-Tdap | 1 | 11y | 18y | NA | NA | NA | - |
| 15 | Non-Tdap | 2 | 0 | 1y | NA | 4w DTaP | 8w Td | 4w Td |
| 16 | Tdap | 2 | 0 | 1y | NC | NC | NC | NC |
| 17 | Non-Tdap | 2 | 1y | 11y | NA | NA | 6m Td | 6m Tdap [b] |
| 18 | Tdap | 2 | 1y | 7y | NC | NC | NC | NC [c] |
| 19 | Tdap | 2 | 7y | 11y | NA | NA | 6m Td | 6m Td [d] |
| 20 | Tdap | 2 | 11y | 18y | NA | NA | NA | 6m Td [e] |
| 21 | Non-Tdap | 2 | 11y | 18y | NA | NA | NA | 6m Tdap [f] |
| 22 | Tdap | 3 | 0 | 7y | - | 0 DTaP | 0 Td | 0 Td |

Not Counted

Not Applicable

**Figure 8**

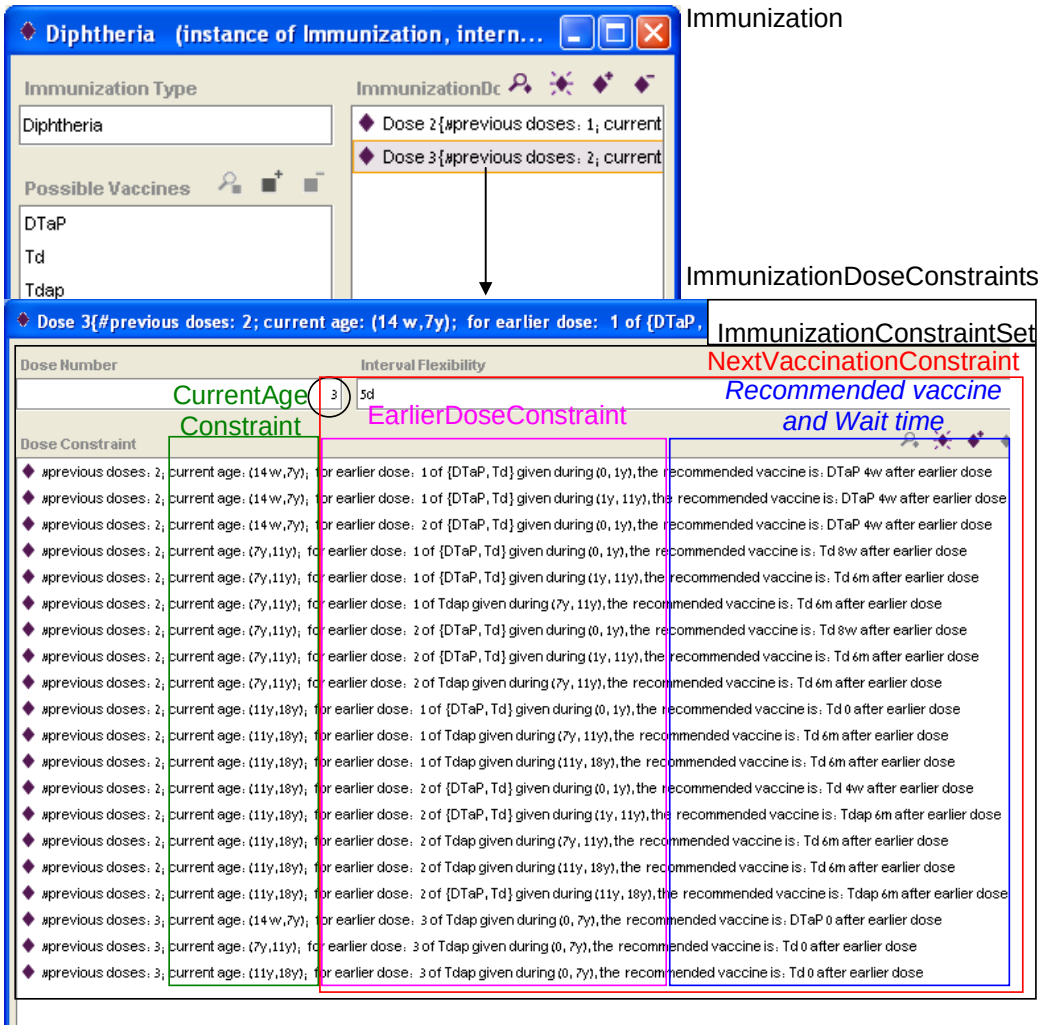


**Figure 9**

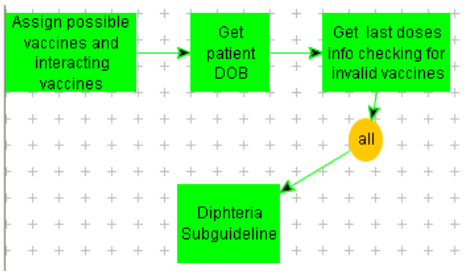


**Figure 10**

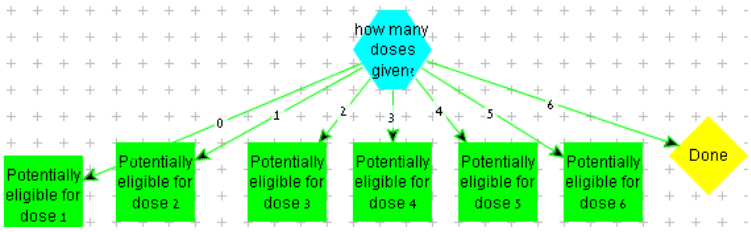


**Figure 11**

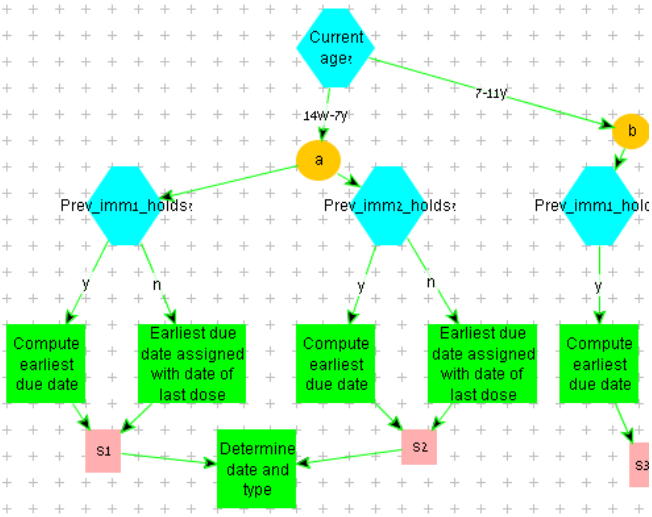


**Figure 12**