# Modeling Contextual Knowledge for Clinical Decision Support

Margarita Sordo, Saverio M. Maviglia, Roberto A. Rocha

Clinical Informatics, Partners eCare, Partners Healthcare, Boston, MA, USA
Brigham and Women's Hospital, Boston, MA, USA
Harvard Medical School, Boston, MA, USA
{msordo, smaviglia, rarocha}@partners.org

**Abstract.** In theory, the logic of decision rules should be atomic. In practice, this is not always possible; initially simple logic statements tend to be overloaded with additional conditions restricting the scope of such rules. By doing so, the original logic soon becomes encumbered with contextual knowledge. Contextual knowledge is re-usable on its own and could be modeled separately from the logic of a rule without losing the intended functionality. We model constraints to explicitly define the context where knowledge of decision rules is actionable. We borrowed concepts from Semantic Web, Complex Adaptive Systems, and Contextual Reasoning. The proposed approach provides the means for identifying and modeling contextual knowledge in a simple, sound manner. The methodology presented herein facilitates rule authoring, fosters consistency in rules implementation and maintenance; facilitates developing authoritative knowledge repositories to promote quality, safety and efficacy of healthcare; and paves the road for future work in knowledge discovery.

## 1 Introduction

One well established best practice for designing rules for decision support systems (DSS) is that "decision rules should be atomic" [1]. Although in theory this is widely agreed upon, in practice this is not always possible. What normally starts as a simple logic statement in the antecedent of a decision rule can develop into what looks as a procedural piece of code when additional conditions are added to the logic in order to restrict the overall rule to a specific context. These additional *constraints* may very well contain relevant information and are perfectly valid, yet from a knowledge representation perspective, they could be considered *contextual knowledge* that could be removed from the logic of the rule. For example, a clinical decision support rule that alerts for an abnormal laboratory test result should specify the laboratory test result in question, a comparison operator and a threshold value. However, the logic of such rule also tends to include conditions for age and gender, among other patient characteristics. This results in having a single rule with a long conditional expression, or multiple variations for the same rule to account for different age groups and genders. This approach creates rules that are less re-usable, with large numbers of pre-coordinated conditions that are difficult to maintain.

Consistent with previous work reported in [2], we believe that decision rules could be modeled as adaptive agents capable of exhibiting specific behavior in response to their environment [3]. This approach allows us to identify and separate contextual knowledge from clinical knowledge and still preserve the intended functionality of decision rules. Furthermore, by clearly separating context from the logic of the rule, we not only facilitate rule implementation and maintenance, but also potentially promote knowledge discovery. The proposed approach to modeling contextual knowledge for decision rules is based on Ontologies, with some basic concepts borrowed from Complex Adaptive Systems and Contextual Reasoning. The work reported herein is part of ongoing efforts to further an authoritative knowledge repository to promote quality, safety and efficacy of healthcare.

### 1.1 Ontologies

Ontologies are a conceptualized and agreed upon collection of entities, relations and instances in a domain of interest where all elements in the ontology are unambiguously described by means of a declarative language and a shared vocabulary [4]. From a modeling perspective, Ontologies are of particular interest for this project as they align with ongoing efforts for extraction, modeling and curation of expert knowledge currently embedded in a variety of systems across Partners Healthcare [5-10]. A large amount of this expert knowledge is in the form of decision rules, and collecting this knowledge from within a variety of disparate systems into a single platform optimized for curation is a crucial strategic initiative at Partners [9].

### 1.2 Complex Adaptive Systems

Holland describes complex adaptive systems (*cas)* as a large number of simple agents, each exhibiting their own behavior in response to external stimuli. There are seven basics (four properties and three mechanisms) central to all *cas* [3]. Our previous work incorporated the four basics (aggregation, tagging, building blocks, and internal models) that we believed best contributed to modeling rule interactions 1. From those four basics, we further explore two mechanisms: a) *Tagging* mechanism facilitates the formation of aggregates and delimitation of boundaries, the latter being a key feature for describing behavior within a context. b) *Building Blocks* mechanism allows reusability and building complex things from repetitions and combinations of simple ones. Our modeling efforts align with both mechanisms in that by tagging a decision rule we successfully remove contextual information from the antecedent of the rule, and explicitly set the boundaries of the environment where such a rule applies. As a result, simplifying the logic of the rule makes it more atomic and reusable.

### 1.3 Contextual Reasoning

Reasoning about context mainly arises from the problem of locality, namely, knowledge that only applies in a restricted world [11]. Such a restricted application might be represented by a set of values that explicitly delimits the boundaries of a world where knowledge is valid. For example, the statement "It's raining" is true only in the context of a given location and time. This is the contextual information we need to explicitly define a more concise representation of the world. There is no need to define all possible aspects of the world that may determine a particular context, only those that are relevant to a given circumstance. Using the above example, we do not need to state whether we are wearing a raincoat or not, or the color of the raincoat for that matter, unless, for some reason such information is relevant. As a result, a *partial* representation – that only includes location and time – suffices for the purposes of setting a context that allows us to determine the validity of the statement "It's raining."

For our modeling purposes we aim at characterizing such a world by borrowing concepts from knowledge representation and reasoning for Ontologies and Semantic Web technologies [12-15]. We specifically base our approach on the metaphor of *context as a box* [16][17]. This metaphor defines contextual metadata in terms of three basic elements: a set of parameters $P_i$ and a value $V_i$ for each parameter $P_i$; a collection of expressions representing the domain at hand; and three abstract forms of reasoning: *expand/contract, push/pop,* and *shifting*, each corresponding to an operation on one of the basic elements of the representation, i.e. parameter, value and expression. Each parameter is seen as a dimension of a *box* (context), which in combination with a value indicate the position of the specific context within a (multi) dimensional space. Hence delimiting what is "inside and outside the *box.*" The remaining of this section describes in more detail these contextual reasoning operations.

*Expand/Contract* operations are based on the intuition that an explicit representation associated with a given context only contains a partial subset of facts – information inside the *box*. As a result, such subset could be *expanded* as more pertinent information is available, or *contracted* if information no longer relevant is removed.

*Push/Pop. Push* operation removes information from inside the *box* and adds it as a parameter value outside the *box*; conversely, *pop* removes a parameter value and adds it as information inside the *box*. For example, if our original statement about the weather were "It's raining in Boston" then we would only have time as a context parameter (the location information is inside the *box*). However, we may choose to remove "Boston" (where we are located) from our statement and *push* it outside the *box* as location information. These operations have a direct effect on both the number of parameters and their values, while delimiting the *box* and the information inside it.

*Shifting* operator allows us to move from one representation to another by changing the value of one or more contextual parameters, as long as the relationship between the parameter's values and the statements inside the *box* is known. A simple example would be to state "Today is raining" in the context where the date/time parameter is set to today's date, and then having the statement "Yesterday it was raining" and *shifting* the value of the date/time parameter to tomorrow's date. We can see that both statements refer to the same fact, though the context was *shifted*.

These operations provide the mechanisms for handling the fundamental aspects of the proposed contextual representation: a) *partiality*, namely the portion of the world being represented; b) *approximation* or level of detail by which the portion of the world is being depicted and; c) the point of view or *perspective* from which the world is being observed. Authors in [16][17] showed that "at a suitable level of abstraction, *a logic of contextual reasoning is precisely a logic of the relationships between partial, approximate and perspectival theories of the world."*

We believe that the *context as a box* metaphor is consistent with the *Tagging* mechanism for *cas* in that both approaches specify a context and delimit the scope of decision rules. Our only departure from the *context as a box* metaphor is that we do not allow any overlapping boxes; all boxes (contexts) must be mutually exclusive, since any overlap may lead to ambiguity, and its inherent difficulties (e.g. limitations of existing rule engines to deal with ambiguity). In the following sections we will further expand on this notion of decision rules as "agents interacting inside a box."

Furthermore, by applying these concepts to *if...then* decision rules we are able to explicitly circumscribe the *context* where actions specified in the *consequent* of a rule are executed when the conditions in the *antecedent* are satisfied.

## 2 Modeling Contextual Knowledge as Constraints

Capturing context is critical for understanding and handling knowledge. This is particularly true in a clinical setting where knowledge embedded in decision rules often times is tailored to specific scenarios. However, it is also most desirable to preserve the generality of rules, ensuring a high degree of reusability and maintainability.

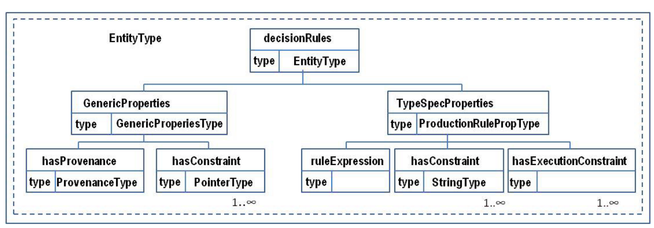


**Fig. 1.** Simplified Schema for Decision Rules.

The schema for decision rules in [2] consists of a) Generic Properties: Provenance [18], and Constraints for the rule; and b) Type Specific Properties, which model the rule expression (described elsewhere 1) and represent the rule itself and its behavior. We currently focus on the type-specific property *hasConstraint*, and we will further describe it in the following section.

### 2.1 Type Specific Property "hasConstraint"

Rule execution can be constrained to narrower scopes by restricting it to more specific contexts. As long as we keep this in mind, we can define as many "sub contexts" as needed for a single rule. For example, an alert for an abnormal laboratory test result may be relevant to all patients regardless of gender and age, and so the constraints in the generic properties of the rule should be set to Gender="ANY", AgeGroup="ANY". The "ANY" value means that such dimension is unrestricted. However, threshold value(s) for the rule may depend on the age of the patient. Therefore, by specifying such threshold values constrained by age groups in the data definition of the rule expression, we can still model an alerting rule as simply as *if* LabResult <operator> threshold *then* alert; where the values assigned to *threshold* are constrained by the context (AgeGroup) where such values apply. Hence, the logic is the same, but the content (operator and threshold values) is defined by the context; in other words, the content – what is inside the box - is dependent on the parameters and values outside the box. This is consistent with the *context as a box* metaphor. Such metaphor and the operations presented in a previous section lay the foundation for our work and allow us to extend/restrict the scope of such rules.

### 2.2 Clinically-Relevant Constraints

In previous research [19-22] authors identified clinically-relevant constraints that should be considered when modeling clinical scenarios. Such constraints are represented by three main context dimensions: a) *Patient*, further subdivided into age, gender, clinical condition, clinical protocol/trial, and health insurance plan; b) *Provider-related* including: group, role, and clinical specialty; and c) *Care Setting* consisting of care setting (inpatient/outpatient), geographic region, facility, department, unit, unit type, room, and bed [21]. From these dimensions, we defined a *partial approximation* by choosing the following four: Patient gender and age group, care setting site, and venue; and abstracted away those dimensions that currently do not advance our purposes. However, as our clinical content evolves, we may choose to further expand such partial representation and include additional dimensions. We modeled the Patient *AgeGroup* dimension as three subcategories: "Prenatal," "Pediatric," and "Adult." Patient *Gender* as "Male" and "Female." Care Setting *Site* currently includes a limited number of Partners Healthcare hospitals: "Brigham and Women's Hospital" (BWH), "Massachusetts General Hospital" (MGH) and, "Newton-Wellesley Hospital" (NW). *Venue* includes "Inpatient," "Outpatient," "Intensive Care Unit" (ICU), and "Emergency Department" (ED). All dimensions support a value of "ANY" to denote that a dimension is explicitly included, but it is left unrestricted. In other words, the scope of a dimension with a value of "ANY" includes all possible values available in such dimension. For example, an *AgeGroup* dimension with a value of "ANY" would include all three age groups, and be satisfied by any of these values.

These four dimensions define a *box* into which we place our decision rules. Values of "ANY" to specify the scope of a dimension allow us to always include all four dimensions, even if not restricted. This has the advantage of providing a fully explicit depiction of the dimensions of our *box*. As shown in Fig. 2a, an unrestricted decision rule is placed in a box with all four dimensions set to a value of "ANY". This means that this hypothetical rule inside the box will apply to any patient regardless of age or gender, at any site or venue.

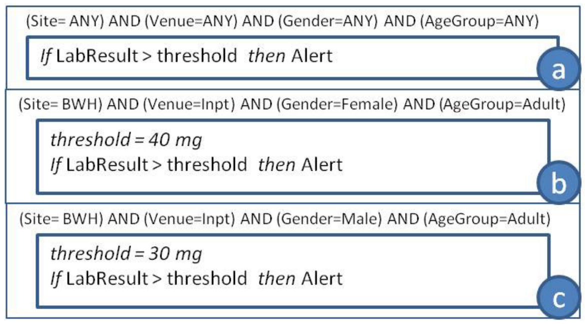


**Fig. 2.** (a) A hypothetical decision rule inside an unrestricted box with dimensions unrestricted with an assigned value of ANY. (b) Rule with context-specific threshold value for adult inpatient females at BWH. (c) Rule with threshold value for adult inpatient males at BWH.

However, if we chose to define a specific threshold value, e.g. 40 mg for adult inpatient women at the BWH, then instead of adding such constraints to the logic of the rule and changing the threshold value to 40 mg, we *push* the constraint values into the parameters (dimensions) outside the box and *expand* (add) the context-specific threshold value inside the box (Fig. 2b). This is equivalent to having both the rule and the threshold value dependent on the context. Additional context-dependent threshold values may exist for a given rule. This would be equivalent to "stacking" boxes. For example, we may define an additional threshold value of 30 mg for adult males for the same site and venue for the above rule. We *shift* the value for the Gender dimension and change the threshold value accordingly (Fig. 2c). This will result in two boxes for the same rule, each having its own context-delimited threshold.

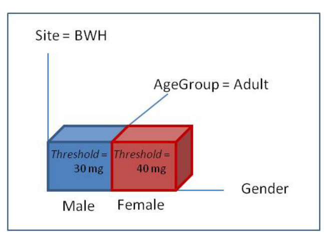


**Fig. 3.** Mutually exclusive boxes stacked to delimit multiple scopes of a single decision rule with context-specific threshold values.

For the sake of graphically representing this notion of "stacking boxes", let us momentarily abstract away (remove) the Venue dimension, and draw a 3D representation (Fig. 3). These two boxes are stacked side-by-side to cover the Gender dimension of the context box. Such (broader) box is divided into two, more restrictive, non-overlapping boxes, each box describing a *partial representation* (portion) of the world, with its own threshold value.

We believe that by having fully circumscribed, mutually exclusive boxes representing partial views of the world we can build a robust contextual representation (consistent with *cas internal models*) of clinical decision rules that can be aggregated and reasoned upon. Similar to the *building blocks* mechanism of *cas*, the rationale behind the *context as a box* approach is to model simple, well-circumscribed behaviors for rules and, then, aggregate them to model more complex behaviors. In the remaining of this paper, we present some exploratory examples and discuss findings and limitations.

## 3 Results

We focused our initial analysis on the clinical content of the Results Manager (RM) computerized application at Partners Healthcare [23]. RM is an application in the outpatient setting consisting of 84 rules that enable clinicians to review, acknowledge, and act upon abnormal/critical results of chemistry, hematology, toxicology, radiology, and cytology tests in a timely manner.

The generic representation of RM alerting rules is as follows: *If* LabResult <*Operator*> <*threshold*> *then* Alert; where LabResult is the value of a given laboratory test result, e.g. “Potassium” (chemistry), “Acetaminophen” (toxicology), “INR” (hematology); *Operator* is a comparison operator to determine whether the laboratory test result is normal/abnormal when compared to a *threshold value. Threshold value* is a reference value for the laboratory test based on values found in the population. Original rules were enhanced to include variations on reference threshold values based on gender, age, and health status of the patient, resulting in a repository of 166 rules. For example, a patient in the ICU or with a chronic condition most likely has some abnormal laboratory test results. This is why, besides age and gender, venue is important to determine the threshold value of a test. We propose the following approach: starting with the general representation of the rule, with all dimensions set to “ANY”, we *push* the constraint on the venue dimension outside the box, and *expand* (add) the context-specific threshold value inside the box. Therefore, in this scenario, the venue would be set to “ICU”, and the threshold value for patients in critical condition would be added inside the box, with the desired effect of having a decision rule with a threshold value targeted to a specific context, that will only trigger an alert under these more delimited circumstances.

Some chemistry rules may need to be restricted to age ranges within an *AgeGroup*. For example, for patients with some suspected thyroid gland dysfunction, a thyrotropin (TSH) chemistry test might be ordered to check the endocrine function of the gland. TSH reference values are age-related, with specific reference values for Pediatric and Adult populations, with further subgroups in the Pediatric population with age ranges of [1 day, 4 days), [4 days, 8days), and [8 days, 18 years). The initial generic representation of this rule is depicted in Fig. 4a. In Fig. 4b we *push* the

Adult *AgeGroup* outside the box and *expand* the operator and threshold value. In Figs. 4c-4e, we *shift* the AgeGroup to Pediatric (PEDI) and include the additional age range, and replace (*contract and expand*) the facts inside the box for threshold and operator.

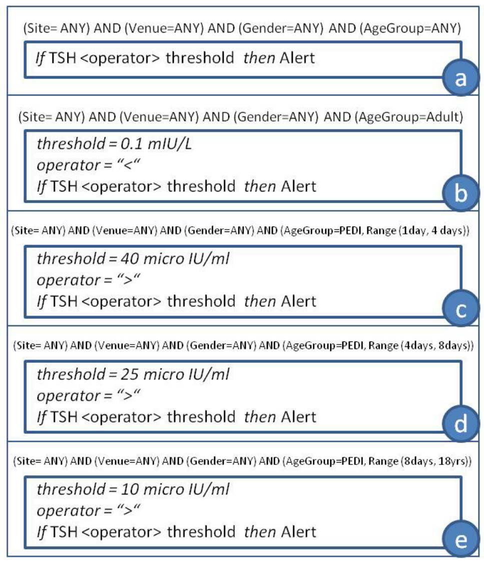


**Fig. 4.** An example of pushing, expanding and shifting parameters for a single rule in multiple contexts.

## 4 Discussion

We have evaluated the feasibility of the modeling strategy proposed above by implementing relevant test scenarios through our analysis of the clinical decision rules in the aforementioned application.

We have shown that with a relatively simple approach we can model abstract contextual knowledge from a variety of rules while preserving the desired functionality. We found that by applying contextual reasoning operators we could expand/contract the context where decision rules apply, while keeping the rule representation as atomic as possible. Further, given that both contextual and clinical knowledge may lie on the same continuum with no fixed point separating them, having the flexibility to vary the degree of approximation of a given representation allows us to regulate the interplay between what goes inside (clinical knowledge) and outside the box (contextual knowledge). By doing so, decision rules can easily be adapted to new contexts, without encumbering the rule logic with extra conditions. We believe this approach is consistent with the internal models and building blocks concepts of *cas*, and it will allow us to model more complex behavior for rules while preserving a simple and sound representation. Likewise, from an ontology perspective, the same mechanisms can be applied to represent both contextual and clinical knowledge as "conceptual building blocks" that can be reason upon while preserving the correctness and expressiveness of the underlying ontology [24].

## 5 Conclusions

Sound and comprehensive approaches are key to accurately modeling knowledge content of any type. In the case of clinical decision rules, it is highly desirable to capture and model not only knowledge pertaining the logic and actions of such rules, but also the context where such knowledge becomes actionable. Currently, contextual knowledge is not identified as such at modeling time, and we strongly believe that it should be. Such contextual knowledge should be removed from the antecedent of a rule, so the logic remains as atomic as possible. The proposed approach provides the means for identifying and modeling contextual knowledge in a simple, yet sound manner. Furthermore, the methodology presented herein facilitates rule authoring, fosters consistency in rule implementation and maintenance; facilitates developing authoritative knowledge repositories to promote quality, safety and efficacy of healthcare; and paves the road for future work in knowledge discovery.

**Acknowledgments.** We would like to thank Beatriz H. Rocha for her valuable expertise and insights that greatly assisted the current research. We also thank Elisa Dell'Oglio, Amanda Fairbanks, Teal Aroy, David Dubois, Karen Bavuso, Sharon Bouyer-Ferullo, Dan Bogaty, Christopher J. Vitale and Priyaranjan Tokachichu for their contributions.

## References

1. Chrisholm M. How to Build a Business Rules Engine. San Francisco: Morgan Kaufmann, 2004.
2. Sordo M, Rocha BH, Morales AA, Maviglia SM, Dell'Oglio E, Fairbanks A, Aroy T, Dubois D, Bouyer-Ferullo S, Rocha RA. Modeling Decision Support Interactions in a Clinical Setting. Stud Health Technol Inform, 2013;192:908-12.
3. Holland JH. Hidden Order: How Adaptation Builds Complexity. Addison-Wesley, 1995.
4. Genesereth MR, Nilsson N. Logical Foundations of Artificial Intelligence. Morgan Kaufmann Publishers: San Mateo, CA., 1987.
5. Greenes RA, Sordo M, Zaccagnini D, Meyer M, Kuperman GJ. Design of a Standards-Based External Rules Engine for Decision Support in a Variety of Application Contexts: Report on a Feasibility Study at Partners Healthcare Systems. Stud Health Technol Inform 2004; 107 (Pt1): 611-615.
6. Sordo M, Hongsermeier T, Kashyap V, Greenes RA. Partners Healthcare Order Set Schema: An Information Model for Management of Clinical Content. In: Yoshida H, Jain A, Ichalkaranje A, Jain LC, Ichalkaranje N, Eds. Advanced Computational Intelligence

Paradigms in Healthcare – 1. Studies in Computational Intelligence Volume 48, 2007; pp. 1-25.

7. Zhou L, Maviglia SM, Goldman DS, Barley A, Lewis J, Hongsermeier T, Rocha RA. A Methodology for the Design, Development and Maintenance of Enterprise Clinical Decision Support Rules. AMIA Annu Symp Proc. 2009 Nov: 1107.
8. Regier R, Gurjar R, Rocha RA. A Clinical rule editor in an Electronic Medical Record Setting: Development, Design and Implementation. AMIA Annu Symp Proc. 2009 Nov: 537.
9. Sordo M. Boxwala A. Grouped Knowledge Elements. In: Greenes, RA, Ed. Clinical Decision Support. The Road to Broad Adoption. 2nd Ed. Academic Press, 2014; pp.499-514.
10. Rocha RA, Maviglia SM, Sordo M, Rocha BH. Clinical Knowledge Management Program. In: Greenes, RA, Ed. Clinical Decision Support. The Road to Broad Adoption. 2nd Ed. Academic Press, 2014; pp.773-818.
11. Giunchiglia, F., 1993, Contextual reasoning. Epistemologia, special issue on I Linguaggi e le Macchine, XVI; pp. 345-364.
12. Ding L, Finin T, Peng Y, Da Silva PP, McGuinness DL. Tracking rdf graph provenance using rdf molecules. Proc Fourth Intl Semantic Web Conference. 2005.
13. Carrol JJ, Bizer C, Hayes P, Stickler P. Named graphs, provenance and trust. Proc 14th Intl Conf World Wide Web. ACM, 2005. pp. 613-622.
14. Homola M, Serafini L, Tamilin A. Modeling Contextualized Knowledge. Proc 2nd Workshop on Context, Information and Ontologies (CIAO). 2010.
15. Contextualized Knowledge Repository. https://dkm.fbk.eu/technologies/contextualized-knowledge-repository; (accessed 09/08/2016).
16. Benerecetti M, Bouquet P, Ghidini C. Contextual Reasoning Distilled. J Expt Theor Artif Intell. 12 (2000). pp. 279-305.
17. Benerecetti M, Bouquet P, Ghidini C. On the Dimensions of Context Dependence. In: Bouquet P, Serafini L, Thomason RH, Eds. Perspectives on Contexts. CSLI Lecture Notes. Center for the Study of Language and Information/SRI, 2007; pp. 1-18.
18. PROV-O: The PROV Ontology. W3C Working Draft. http://www.w3.org/TR/prov-o/; (accessed 09/02/2016).
19. Maviglia SM, Yoon CS, Bates DW, Kuperman G. KnowledgeLink: impact of context-sensitive information retrieval on clinicians' information needs. J Am Med Inform Assoc, 13(1); pp. 67-73.
20. Rocha RA, Maviglia SM, Sordo M, Rocha BH. Clinical Knowledge Management Program. In: Greenes, RA, Ed. Clinical Decision Support. The Road to Broad Adoption. 2nd Ed. Academic Press, 2014; pp.773-818.
21. Del Fiol G, Rocha RA, Bradshaw RL, Hulse NC, Roemer LK. An XML model that enables the development of complex order sets by clinical experts. IEEE Trans Inf Technol Biomed 2005; 9, 216-28.
22. Del Fiol G, Huse V, Strasberg HR, Maviglia SM, Curtis C, Cimino JJ. Implementations of the HL7 Context-Aware Knowledge Retrieval (“Infobutton”) Standard: challenges, strengths, limitations, and uptake. J Biomed Inform. 2012 Aug; 45(4); 726-35.
23. Poon EG, Wang SJ, Gandhi TK, Bates DW, Kuperman GJ. Design and implementation of a comprehensive outpatient Results Manager. J Biomed Inform. 2003; 36(1-2):80-91.
24. Rector A, Rogers J, Roberts A, Wroe C. Scale and context: issues in ontologies to link health- and bio-informatics. AMIA Annu Symp Proc. 2002: 642-6.