

# Analysis of Clinical Record Data for Anticoagulation Management within an EHR System

T. Austin<sup>\*,1</sup>, D. Kalra<sup>1</sup>, N.C. Lea<sup>1</sup>, D.L.H. Patterson<sup>2</sup> and D. Ingram<sup>1</sup>

<sup>1</sup>Centre for Health Informatics and Multiprofessional Education (CHIME), University College London (UCL), UK

<sup>2</sup>Whittington Hospital, Archway, London, UK

**Abstract:** *Objectives:* This paper reports an evaluation of the properties of a generic electronic health record information model that were actually required and used when importing an existing clinical application into a generic EHR repository.

*Method:* A generic EHR repository and system were developed as part of the EU Projects Synapses and SynEx. A Web application to support the management of anticoagulation therapy was developed to interface to the EHR system, and deployed within a north London hospital with five years of cumulative clinical data from the previous existing anticoagulation management application. This offered the opportunity to critique those parts of the generic EHR that were actually needed to represent the legacy data.

*Results:* The anticoagulation records from 3,226 patients were imported and represented using over 900,000 Record Components (i.e. each patient's record contained on average 289 nodes), of which around two thirds were Element Items (i.e. value-containing leaf nodes), the remainder being container nodes (i.e. headings and sub-headings). Each node is capable of incorporating a rich set of context properties, but in reality it was found that many properties were not used at all, and some infrequently (e.g. only around 0.5% of Record Components had ever been revised).

*Conclusions:* The process of developing generic EHR information models, arising from research and embodied within new-generation interoperability standards and specifications, has been strongly driven by requirements. These requirements have been gathered primarily by collecting use cases and examples from clinical communities, and been added to successive generations of these models. A priority setting approach has not to date been pursued - all requirements have been received and almost invariably met. This work has shown how little of the resulting model is actually needed to represent useful and usable clinical data. A wider range of such evaluations, looking at different kinds of existing clinical system, is needed to balance the theoretical requirements gathering processes, in order to result in EHR information models of an ideal level of complexity.

## INTRODUCTION

In order to practice high quality and safe medicine, healthcare professionals need access to patient information that is as complete as possible, available at any relevant point of care and within a suitable time. The basis of this is the health record. This is not necessarily a single integrated record but rather consists of many separate independent autonomous components distributed throughout the health sector. In order for health professionals to share care, these separate components must be brought together and integrated whenever needed. Obstacles to achieving this with paper records are well known, and the electronic health record (EHR) is a much-anticipated solution to the challenge, offering the possibility of real-time seamless information flows to any point of care. In order to achieve such a vision, the European Commission has funded several research and development projects over the past fifteen years, aimed at investigating the clinical, ethico-legal and

technical requirements for an internationally-interoperable EHR and at prototyping information architectures for it including GEHR [1], Synapses [2], and SynEx [3].

The Synapses model provided the most significant underpinning to the 1999 European (CEN) EHR Communications pre-standard ENV13606 [4], and the UCL EHR server is one of a limited number of live implementations that also conform to that pre-standard. This research background has strongly influenced more recent generic EHR information and interoperability models such as ISO/EN 13606 [5]. Few evaluations have been published about the experience of implementing generic EHR information models, and those have generally focussed on the ease of implementation or on mapping a clinical application to the model. No previous studies, to the knowledge of the authors, have been published that review the data content of an EHR server. Whilst the models meet a large number of reported clinical and ethico-legal requirements, concerns have been expressed informally by industry that they might be unnecessarily complex.

This paper reports an evaluation of the properties of a generic electronic health record information model that were actually required and used when importing an existing clinical application into a generic EHR repository.

\*Address correspondence to this author at the Centre for Health Informatics and Multiprofessional Education (CHIME), UCL, 4th Floor, Holborn Union Building, Highgate Hill, London, N19 5LW, UK; Tel: +44 20 7288 3372; Fax: +44 20 7722 3322; E-mail: t.austin@chime.ucl.ac.uk

## METHOD

The Synapses project, building on the previously published Good European Health Record Architecture, investigated user requirements for a generic federated (distributed) EHR architecture, and from this developed an information model to represent the generic properties of EHR data including those necessary to capture clinical and medico-legal context. A generic EHR repository and middleware services were developed at UCL as part of these projects. A Web application to support the management of anticoagulation therapy was developed to interface to the EHR system, and deployed within a north London hospital [6]. Several years of cumulative clinical data from the previous existing anticoagulation management application were imported into the EHR as part of this process.

### The Record Model

In order to represent record data in a scalable and extensible way Synapses pioneered a two-level modelling approach [7,8]. This distinguished the generic identification and medico-legal hierarchy constructs of the EHR model (known as the Synapses Object Model, or *SynOM*) from a formalism to represent and share domain knowledge about specific clinical data structures used within the EHR (the Synapses Object Dictionary, or *SynOD*). This enables the clinical structures needed for recording EHR data to evolve without impact upon computer systems that can already persist records. This dual modelling approach has since been adopted more widely internationally in *openEHR* [9,10], CEN [4], and HL7 [11,12].

Fig. (1) shows the principal components of the Synapses federated health record model (*SynOM*), which is nearly identical to ENV13606:1999. In this figure, the concrete classes are shown in a shaded colour, and the abstract classes are in white.

#### RecordComponent

*RecordComponent* is the abstract base class for *RecordItemComplex* and *RecordItem*. The complete set of attributes and their data types broadly define the commonality applicable to all major classes of the *SynOM* for: record authorship; ownership and duty of care responsibilities; the subject of care; dates and times of healthcare actions and of their recording; version control; access rights; and to a limited extent, emphasis and presentation.

The *SynOM* distinguishes between the aggregation necessary to convey compound clinical concepts and the aggregation within a record that groups observations relating to the healthcare activities performed. An example of the former would be blood pressure, which is a compound concept composed of systolic and diastolic values. An example of the latter would be the grouping together of observations under a general heading of Physical Examination. The *RecordItem* and *RecordItemComplex* constructs respectively represent these two categories of aggregation.

#### RecordItemComplex (RIC)

Two broad categories of RIC are defined in the *SynOM* through two abstract sub-classes of *RecordItemComplex*.

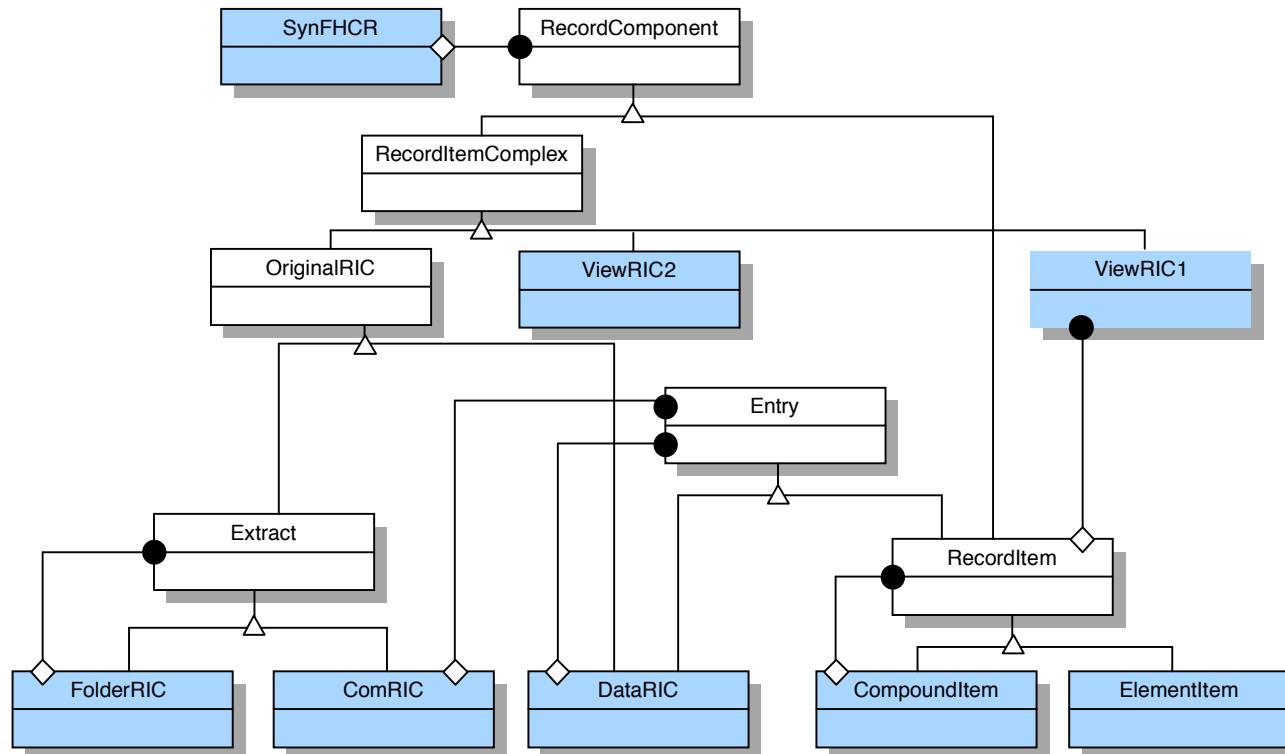


Fig. (1). The synapses *SynOM* inheritance diagram (from [8]).

1. OriginalRIC: this set of classes represents the original organisational structure (grouping) of sets of record entries as defined by the author(s) of those entries; it provides the medico-legal representation of the underlying information.
2. ViewRIC1 and ViewRIC2: this set of classes provide the means by which alternative groupings and subsets of the original information may be organised and preserved as permanent views in a patient's record, unlike those generic views provided in an ad hoc way by a client system. A ViewRIC1 is derived through the use of a predefined query procedure and a ViewRIC2 provides a static view of original information through a set of references to the original entries or groups of entries.

### **OriginalRIC**

Three concrete classes of OriginalRIC are defined in the SynOM to provide for the nested aggregation of original groupings for record entries.

1. FolderRIC: FolderRICs define the highest levels of organisation within healthcare records. They will often be used to group large sets of record entries within departments or sites, over periods of time, or to demarcate a prolonged illness and its treatment. Examples of FolderRICs include an episode of care, an inpatient stay, or one stage of a disease process. FolderRICs can contain other FolderRICs, and/or ComRICs.
2. ComRIC: A medico-legal set of record entries required by the author to be kept together to preserve meaning when information is physically moved or copied to another persistent store. This is to ensure that all persistent EHR stores comprise whole ComRICs. The ComRIC defines the medico-legal cohort for the inclusion of new entries within an EHR: any new EHR entry (even if stored on a local feeder) must be a whole ComRIC. ComRICs cannot contain other ComRICs or FolderRICs.
3. DataRIC: This class is intended for grouping observations under headings within a ComRIC. It therefore provides for fine granularity grouping and labelling of record entries with names that relate the clinical concepts to the healthcare activities and processes surrounding the patient. DataRICs may contain other DataRICs and/or RecordItems. They cannot contain ComRICs or FolderRICs.

### **RecordItem**

This class defines the structure of the individual clinical entries within a record. It was defined in Synapses as the smallest unit of information which remains meaningful as an entry in its own right. This abstract class provides an aggregation construct for clinical concepts that are then composed of one or more individual named clinical values (e.g. pulse or blood pressure) using the concrete classes CompoundItem and ElementItem. These entries may be aggregated within a hierarchy to represent complex clinical concepts but such a composition is distinct from the record structure grouping hierarchy provided by the RecordItemComplex classes.

The ElementItem supports several data types for the content value that may be assigned to any element entry. These include individual or multiple numeric quantities, dates, images and multimedia, plain textual elements or clinical terms, and references to people and things. The RecordItem class also inherits the medico-legal attributes defined in the RecordComponent class.

The aggregation relationship between the RecordItem and other classes in the SynOM is given in Fig. (2).

The attributes of the RecordComponent class are inherited throughout the SynOM class hierarchy and may acquire instance values at any level of the record. This is so that the greatest number of legacy systems may be directly represented within the object model. Some of these attributes have been defined as mandatory and must be incorporated within any EHR managed by a Synapses server in order to comply with the specification. If mandatory information is not present in an underlying feeder system's data, then a value indicative of this must be included within the RecordComponent object. Other attributes, defined as optional, have been included as recommendations for good practice. A summary of the attributes of RecordComponent is given below.

**Class Attributes:** The class name, its unique identifier, the identifier of the Object Dictionary (SynOD) entry that defines its semantics, and an optional identifier for the original parent.

**Data Subjects:** an identifier for the subject of care and optionally for the information subject if not the subject of care.

**Authorship, Ownership and Duty of Care:** The identities of the parties providing, recording, and revising the EHR data, the parties clinically responsible and legally responsible for care, and the original EHR system from which data were acquired (if they originated elsewhere).

**Details of Actions and their Recording:** When and where health care activities took place, when they actually occurred and when they were recorded.

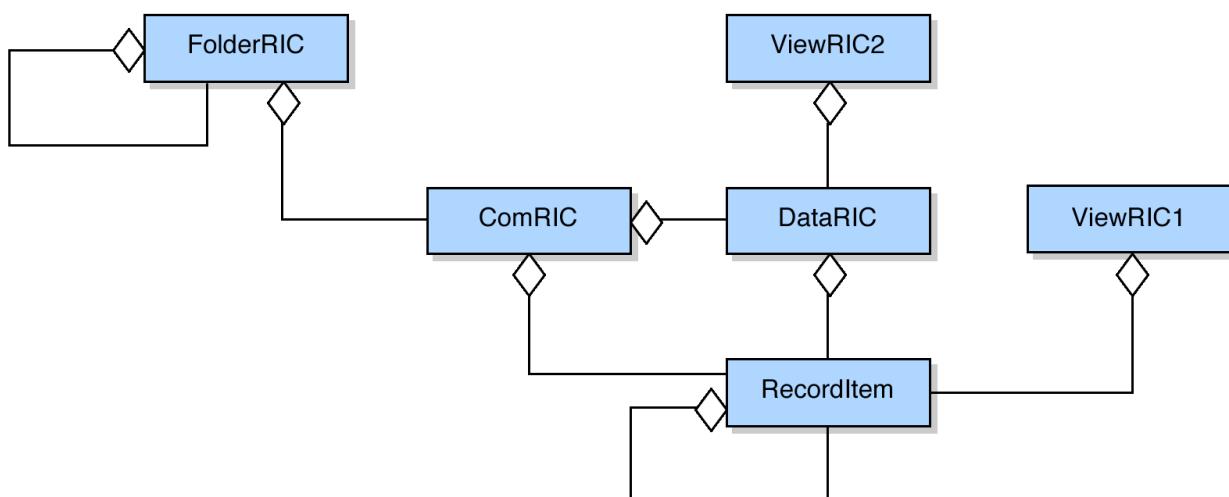
**Version Control:** References that link former and current versions of the EHR data, and the attestation status of each version.

**Access Control:** A set of values reflecting an ordered set of sensitivity levels.

**Miscellaneous:** Optional properties to mark data as emphasised, to add in a free-text author's comment, and to represent links between this and any other parts of the EHR to show the clinical relationships between entries.

### **The SynOD**

A complement to the generic EHR model described above is the Synapses Object Dictionary (the SynOD) that defines the hierarchy of clinical entities that any given EHR server must recognise and manage within its repository. This is effectively an EHR domain model for the domains in which the server is to be deployed. The SynOD defines a set of constraints that ensure its contained entities are consistently represented, and an updated version now features in ISO EN13606 as the "archetype". An example



**Fig. (2).** Aggregation structure of the SynOM (from [2]).

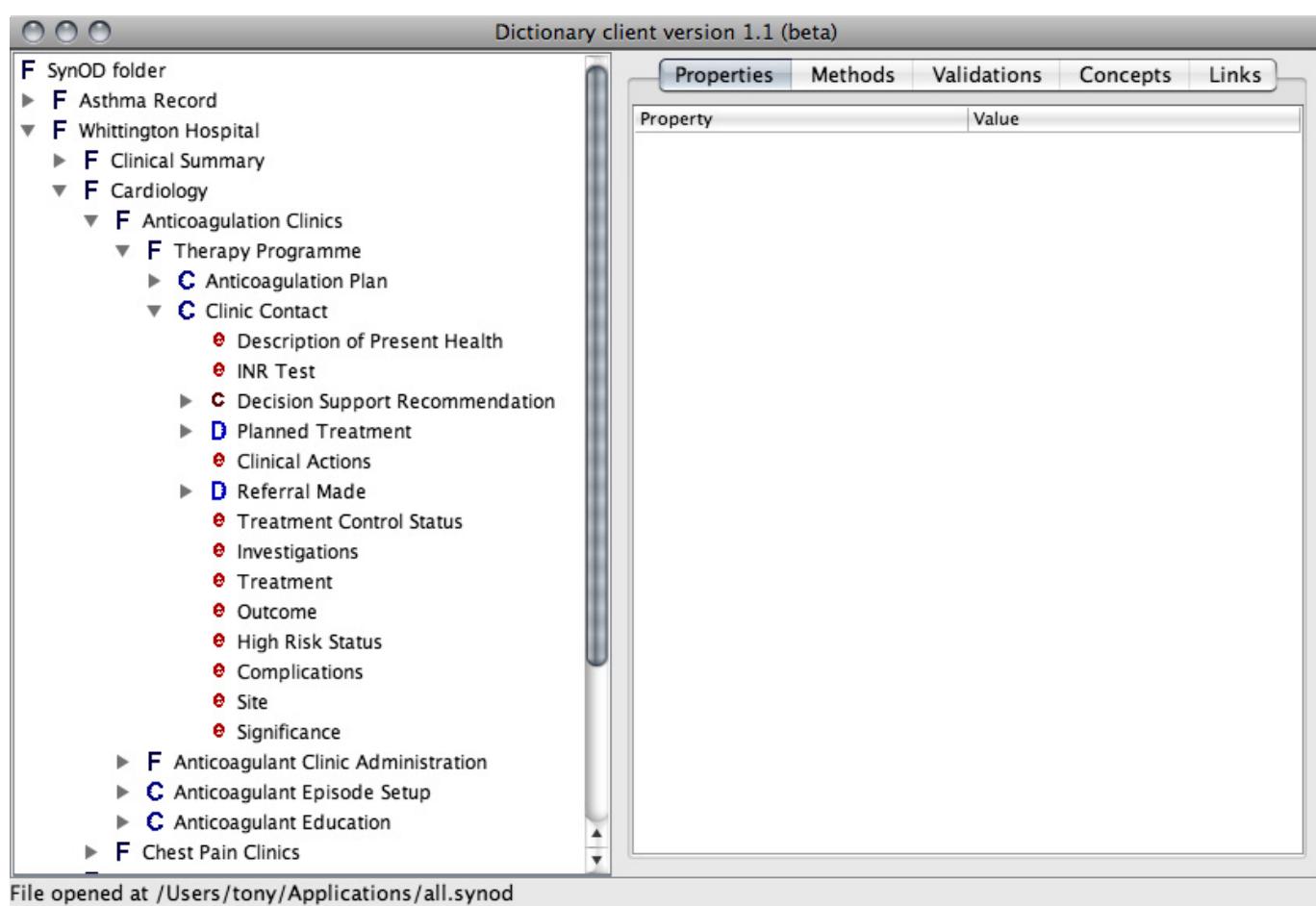
portion of the SynOD defined for the anticoagulation system is shown in Fig. (3) below.

### The Practical Context

The Department of Cardiovascular Medicine at the Whittington Hospital in London is a busy department where most facets of cardiac care are administered. It provides for approximately 1,800 emergency inpatient admissions, 7,000

outpatients and 20,000 cardiac investigations per annum. It runs several outpatient clinics and partners closely with GP practices in North London [6].

Warfarin is a drug that thins the blood and prevents emboli or clots passing from the heart and causing such problems as strokes, or passing to the heart lumps from the veins and causing a pulmonary embolus. The dose of



**Fig. (3).** The SynOD model of an anticoagulation clinic contact.

warfarin has to be carefully controlled using a blood test which needs to be maintained in a normal therapeutic range. Poor warfarin control leads to increased morbidity and mortality, for example through abnormal bleeding problems, and longer hospital stays. Computer based decision support can provide valuable assistance to the clinician and improve patient care. Chin [13] found that decision support can improve adherence to clinical guidelines and further positive outcomes include rapid response to critical laboratory results and adverse drug reactions [14-16].

Vadher [17] showed that warfarin control is an area of cardiovascular medicine that can be automated both in terms of the recommendation of an appropriate dosage and in the regulation of clinic attendance. He developed an algorithm for warfarin dosing which was evaluated in a randomised controlled trial at the Whittington Hospital [18,19]. It was found that using the decision support tool the median time to achieve a stable dose was significantly lower, and both inpatients and outpatients spent longer within the therapeutic range than did patients dosed by unassisted trainee doctors. These benefits can be realised by nurse practitioners with appropriate training and backed by a computerised clinical application and advisory system [20].

### **The Anticoagulant Application**

When patients are referred to the anticoagulant clinic, as an inpatient from the ward or as an outpatient by the patient's GP, some basic data about the background clinical history, indications for anticoagulation, current problems, allergies and medication are recorded. A treatment plan which includes a target level of anticoagulation and a planned duration, is defined. Each clinic visit at which the patient has a blood test to measure the current level of anticoagulant control, results in a fresh appraisal of the correct dose of warfarin to be taken and of the next appropriate review date. These clinical decisions are supported by a computerised decision support system implemented as a Java<sup>TM</sup> middleware component using the algorithm outlined above.

This study was conducted in the context of a migration from a legacy warfarin control application written in Microsoft<sup>®</sup> Access<sup>TM</sup> 2.0, to a new version also written in Java and running as a Web application. An example of the screens provided by the new system is given in Fig. (4). This is the Clinic Contact screen, the main one used by clinic staff when dosing patients on warfarin. The blood test used to control the patient in a therapeutic range is called the International Normalised Ratio or INR. A detailed breakdown of the most recent such result is given at the top of this screen, with summary results of the previous five readings (where these are available) given in a table at the bottom. Decision support (abbreviated "DS" in the figure) is provided to recommend a next dose and next visit date for those patients within a therapeutic range.

### **RESULTS**

The EHR data created in the Cardiology Department derived from one year of using the migrated anticoagulant system described in this paper and about 4 years of using the Microsoft Access predecessor version, resulting in a total of

five years worth of EHR data within the server. The definitions in the SynOD were created to correspond most effectively to the legacy database and to meet contemporary clinical requirements for the system. By examining the properties of the instances over the entire period it is possible to consider whether the richness of the SynOM and SynOD were fully exploited with this domain architecture.

### **Record Component Analyses**

The five years of data comprised 932,606 RecordComponents in the system in total and of these, 638,199 were ElementItems. More information is shown in Fig. (5). It is not possible to plot the change in these data sets over time since the preceding system did not capture the date of recording. Fig. (6) shows the way in which the ElementItem content is broken down by its value's data type. Slightly more than two thirds of the ElementItems (435,936 out of 638,199) were a type of Numeric quantity. There were very similar amounts of date-related and textual information.

There are no multimedia ("Bulky") data items in this application's record set. The only candidate for such data within the application is related to patient education and medical student teaching. However a simple Web file repository could easily satisfy this requirement (since the resources are not patient-specific) and so no formal entities for multimedia data were defined in this SynOD. The very small numbers of "PersonDevice" elements are used to refer to healthcare practitioners that are outside the current episode of care. The GP is occasionally mentioned as the "referring healthcare professional". Since the RecordComponent has a very rich set of attributes for describing personnel directly connected to the episode (such as the recording and legally responsible agents of care) the need for an explicit content class to hold this sort of information might be called into question based on this data. This is especially true as there is an opportunity to attach a person or device identifier to any Numeric value as well.

Fig. (7) shows how the set of Numeric instances breaks down. The majority of these are simple quantities with single values. All single values are stored internally as double precision floating point numbers so the difference between integer and real values is not reflected in the database. A special case of the single value, the Boolean, represents the next largest category. Together these two single-value storage options account for 99% of the Numeric instances in the database.

Only 4,020 of the total 435,936 Numeric items in the database are modelled as ranges. These are the "Target INR" ranges that are included in each created anticoagulant treatment plan. There are no uses made in the database, of ratios. There are two candidate occasions where ratios might have been used; to represent the units of a value, and for clinical compounds like blood pressure that are expressed in ratio form (e.g. 120/80). However a "units" item can be explicitly attached to a Numeric quantity so that a ratio to contain them isn't required. The second case is modelled as a compound of two definitions in the SynOD. Further implementation experience is required to establish the need for explicit modelling of ratios, based on these data.

Whittington Hospital Cardiology Department Anti-Coagulant Client - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address http://quandary.chime.ud.ac.uk/anticoag/servlet/acClinics?patID=whCardioAcPatients-ecf09f834c\_0&index=1 Go Links

Google Search Web Search Site PageRank Page Info Up Highlight

[Search](#) [Demographics](#) [Plans](#) Clinics [Episodes](#) [Patient's Summary](#) [Clinician's Summary](#)

 Whittington Hospital Cardiology Department Anti-Coagulant Client

You requested information on patient:  
Peter Smith (12 December 1912)

New Clinic Contact?

Date:	01/04/2002 08.36
Present Health:	~

Target INR	INR Test	Treatment Control Status	Needs Clinic Visit?	Planned End Date
2.5	2.5	Maintenance	No	01 May 2002

Present Health

DS Dose	Accepted?	DS Interval	Accepted?
1.0	Yes	14.0	Yes
Override Dose	Override Interval	Next Appointment	Days of Anticoag. Control
~	~	15/04/2002 08.36	~

Decision Support & Treatment

Authors Comment (Clinical):	Patient is compliant with treatment
Patient Instructions:	Keep it up!
*** Edit ***	

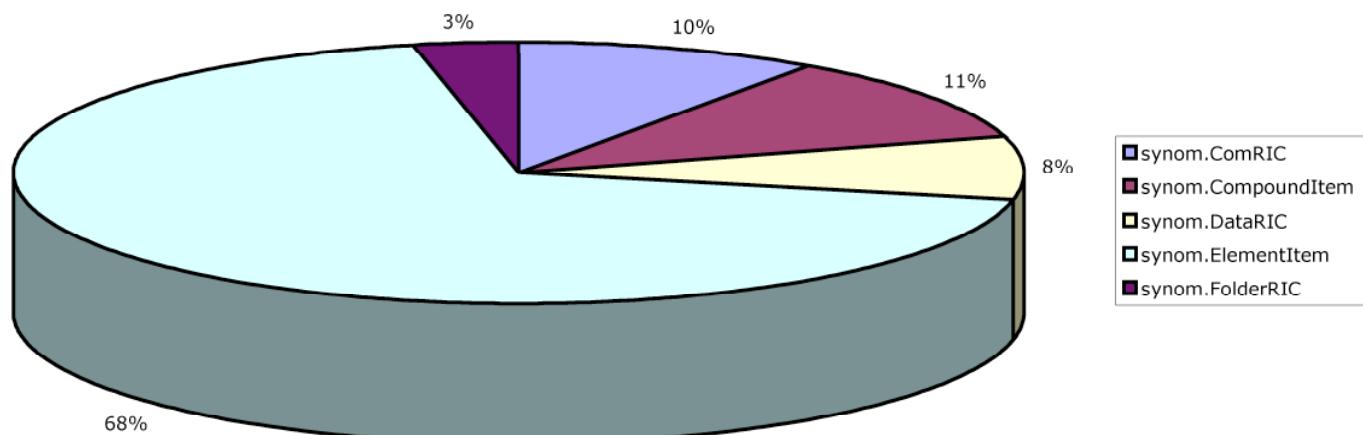
[Complications](#) [Referrals](#)  
[Previous](#) [Next](#)  
[1](#) [2](#) [3](#)

Recent Past Readings

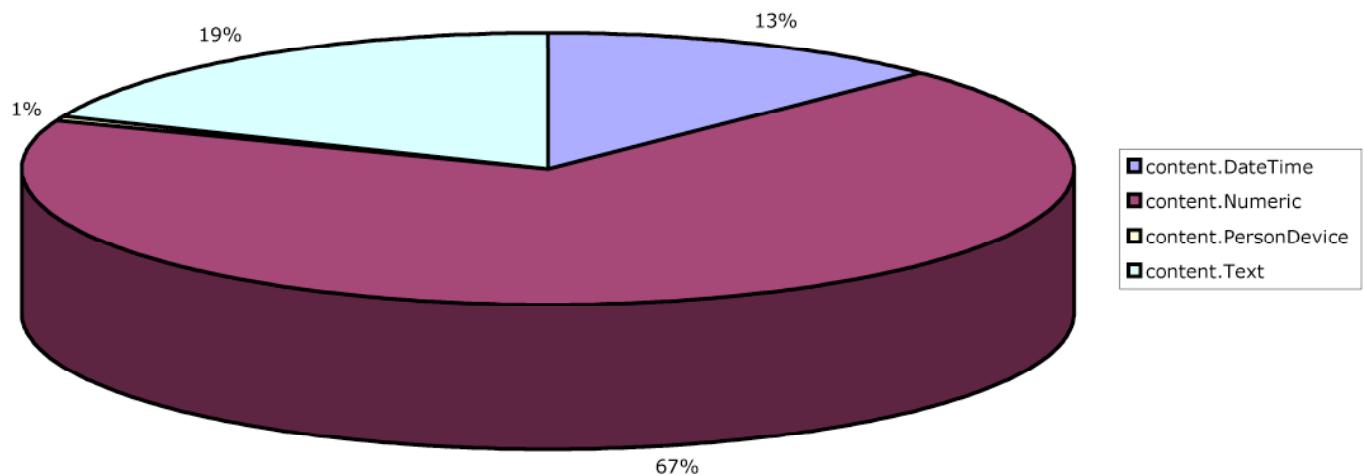
Date	INR Test	DS Dose	Accepted ?	DS Interval	Accepted ?	Override Dose	Override Interval
01 March 2002	2.5	~	~	~	~	1.0	1.0

Done Internet

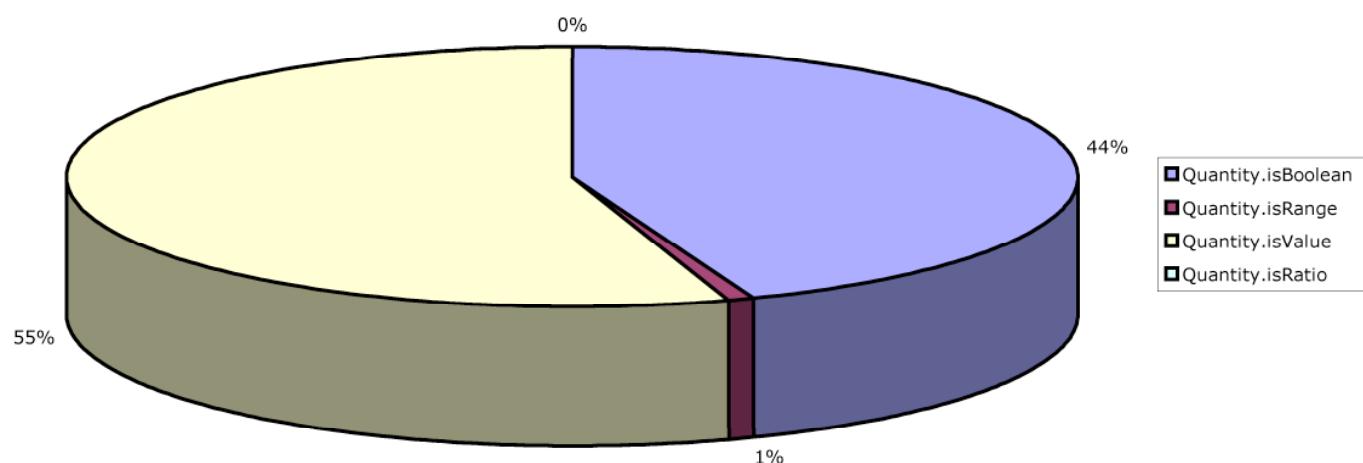
Fig. (4). The anticoagulant application clinic contact screen.



**Fig. (5).** Breakdown of RecordComponent type.



**Fig. (6).** Breakdown of ContentClass type.



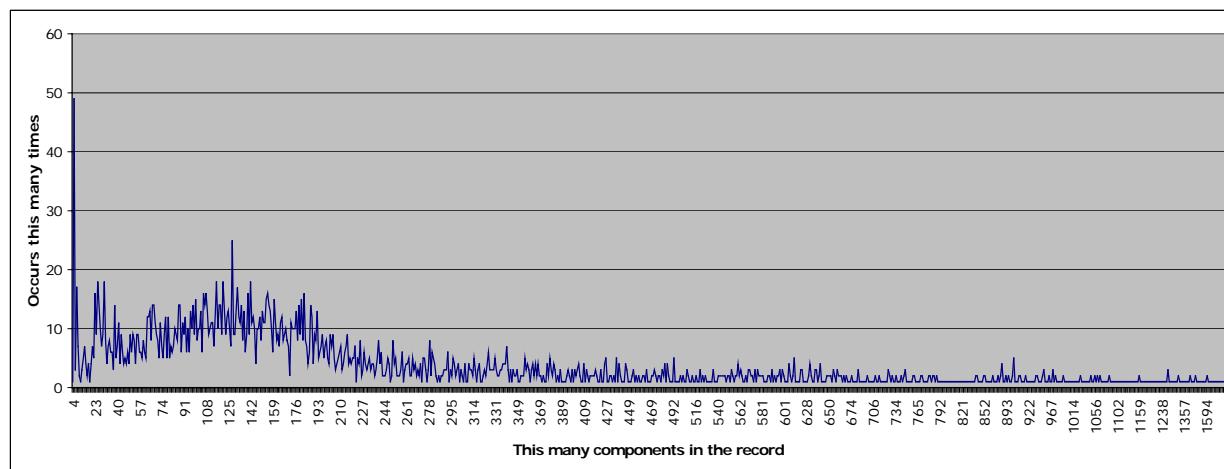
**Fig. (7).** Breakdown of numeric value usage.

### Patient Records Analyses

Within the combined legacy and recent data there are 3,226 patients. In the previous system patients managed in the community and within the hospital were kept in separate databases and reconciled together on a monthly basis (occasionally sooner if the situation demanded). This is now no longer necessary since the hospital and community

patients are managed through the common Web interface presented by the application.

The least number of RecordComponents within the database comprising a single record is 4. The largest number of components is 3,223. The average record size is about 289 components per record with a median value of 173. The distribution of RecordComponent quantities is given in Fig. (8).



**Fig. (8).** Number of components per record.

The SynOD entity appearing least in the database (“Anticoagulation Precautions”) occurs only 5 times. The next lowest (“Outstanding Issues”) occurs only 8 times. Both of these components pertain to patient education and a conclusion might be that the opportunity to educate the patient is not being taken or recorded in the context of a busy clinic. Certain definitions (such as the top-level container “Whittington Hospital”) are designed to be included only once in the patient record and do indeed occur 3,226 times in the database as a whole. 55 of the set of 85 object definitions (just over 64%) occur on average less than once per record. This would suggest that many of the entities have been added on the basis of clinical expectation of use rather than known use, and is a finding echoed anecdotally by other clinical application developers: that the expert users who help define a system are more enthusiastic about the data that “should” be captured than most of their colleagues prove to be in actually capturing it.

The process of providing anticoagulant therapy control is one of regular visits to a clinic to have an INR test and to have the dose of warfarin reaffirmed or adjusted. It is therefore not surprising that “Clinic Contact” represents the highest proportion of RecordComponent instances at 71,496 (just over 22 times per patient record on average). It is perhaps more surprising that slightly fewer INR Tests occur (66,106 or an average of just over 20 times per record). That means around 10% of Clinic Contacts seem not to include an INR test result. Next appointments (“Next Appointment Date”) are made 68,604 times, or for almost 96% of Clinic Contacts. The other 4% presumably represent patients that are discharged.

Decision support for dosing (“DS Dose”) is given 46,467 times and for interval calculations (“DS Interval”) slightly higher at 51,321. These are 65% and 71% of the total Clinic Contacts respectively. There is interest within the department to expand the number of therapeutic ranges on which the decision support engine can advise and there is evidence here that a substantial proportion of patients are presenting with dosing requirements that are not met by the present decision support algorithms.

### Structural Attributes

Each RecordComponent retains a considerable and highly expressive set of attributes that can describe contextual information about the instantiation of that instance, as described earlier in this paper. These are actually fields in the Java class for RecordComponent. A significant number of “blank spaces” in the attribute list would lead to a conclusion that an alternative representation for attributes should be sought. For example they might be better stored as a table mapping attribute names to values or as a linked list.

Certain attributes have been treated as read-only in the anticoagulant system. The “rights” attribute giving a first level of security protection within the record is never modified from its default value. This represents the right of the patient’s own clinician to read the record and offer care but does not permit audit or administrative staff.

No item in the anticoagulant record is marked as being emphasised although it is worth noting that extensive use was made of the Emphasis attribute in another project which focussed on Asthma (Medicate). This used the attribute to point out which were the unusual peak flow and symptom readings worthy of further investigation. In that project, coloured text boxes were used on Web forms to highlight those items thus marked in the record.

No HealthcareActivity or ObservationDateTimes have been used. RecordingDateTime is added comprehensively and there are record entries with a DateTime content type defined in the SynOD (most notably the “Next Appointment Date”) but little has been made of the other dates provided for as attributes.

Three types of healthcare provider may be described; the Recording, the Responsible and the LegallyResponsible Agent. The first of these is not used since the typical pattern in the department is not one of administrative personnel entering data for a clinical team member to sign off. Clinic Contacts do exploit the ResponsibleHealthcareAgent field but only 29,265 of 71,496 total Clinic Contacts (40%) use this feature. This proportion is consistent with the higher level of clinical activity in the most recent (fifth) year of this study, and it is therefore probable that it is now being used in

the new Web application but was not recorded in the legacy data. The LegallyResponsibleHealthcareAgent and 6 other attributes, the RecordingLocale, the HealthcareActivity Location, the InformationProvider, the EHCRSource, the AuthorisationStatus and the SubjectOfInformation are all recorded automatically by the record server mechanics. Generally speaking, these have all been inserted into the database in exactly the same proportion so it is likely that these data have also only been collected since the new record server became available.

Of the 932,606 RecordComponents in the database, 29,140 (9 RecordComponents per patient on average) have an AuthorsComment attached. When developing the model this field was seen as potentially problematic since it makes it possible for a significant amount of “record information” to exist in the comment fields (orthogonally to the structure defined in the SynOD). These would have to be separately checked in order to derive the full meaning of the record. However, this complexity was contained by limiting user access to the author’s comment attribute *via* the Web applications and the predominant use of the field here is to leave instructions for the patient or to explain any override of the decision support.

5,537 of the RecordComponents have their “Revised By” flag set, indicating that another version of them exists. However the capacity to record revisions has only been provided since the new Web application was created so this represents not 1 or 2 records per patient over the lifetime of their care, but over a single year. Fig. (9) shows how the revisions are broken down by SynOD class.

## DISCUSSION

The results demonstrate some of the ways in which healthcare data created by clinical personnel is represented within a generic EHR.

The results naturally reflect the domain of interest: cardiovascular outpatient care, and in particular anticoagulation management. Inevitably when considering any patient as a whole, the information relevant to one clinical domain is part of a broader virtual EHR. Data integration from these other systems, such as GP systems, was not feasible in this demonstrator but it is recognised that the inclusion of such data may well have led to a more complete coverage of the attribute set or aggregation classes. Since enterprises legally responsible for maintenance of a portion of a patient’s data cannot easily merge their holdings into one single “super-database” this reality is precluded in healthcare for the time being, pending the maturation of national e-health infrastructures.

## Hierarchical Structure

The number of RecordComponents that are not ElementItems (almost a third, 32%) justifies the hypothesis that record data is hierarchical in nature. If this were not the case far fewer entries in the record database providing purely heading and sub-heading context would be expected. The construction of the record model inevitably requires a certain level of aggregation that cannot be avoided, but if the observed instances were only artefacts of the mapping process there would be fewer of them. Certainly not one third of the total output.

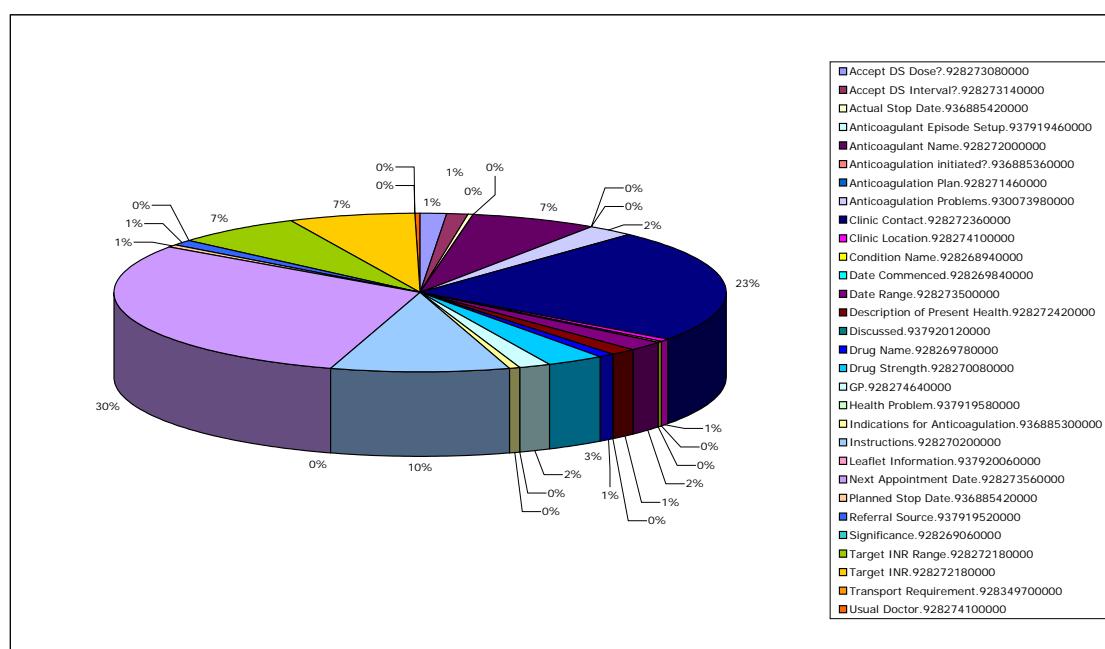


Fig. (9). Revision breakdown.

A more complex variant of the argument is that since in-house developers were responsible for the Web application front-end it would have been possible for them to skew the results in favour of a highly aggregated record structure for the sake of proving that one is needed. This is a much more difficult claim to refute but the team has endeavoured to maintain a separation between developments on back-end systems and those on the Web applications themselves. The same people do not generally work on both.

It is true that those record structures over which there was greater control (e.g. where there had not been a legacy feeder system to migrate) resulted in cleaner, flatter structures. This does not necessarily result in fewer contextual entries however, since the medico-legal requirement to commit a cohort of data in a ComRIC remains. These results, in particular the breakdown depicted in Fig. (5), indicate that the EHR hierarchy classes of Folder, Com, Data, Compound and Element all have value in organising the data within an EHR.

#### **Attribute Usage**

It is reasonable to assess what developers found of value in the model during the construction of the clinical application, since this is how the server software will be predominantly used. The decisions on what to record and where, are in almost all real cases taken by groups developing clinical applications in partnership with their clinical users.

Usage of the attributes inherent in RecordComponents and ContentClasses has been so low in comparison to the number of components being committed that it might be concluded there is no requirement for them to be inherent in the objects forming the record at all. The inclusion of key dates and parties within a particular kind of entry is so specific that these would appear to be better included as explicit entities within the SynOD rather than as generic properties of all RecordComponents.

The results suggest that:

- the record model should probably only include a few specific DateTime attributes (for when the data were clinically acquired and when they were recorded), and that other dates and times should be explicitly labelled within the SynOD;
- the Author'sComment option is rarely used if the record is appropriately structured, and should not be part of the generic record model; if needed specifically-labelled commenting leaf nodes should be included within relevant SynOD definitions;
- it is less important for the model to represent the roles of all healthcare actors and their locations explicitly;
- attributes to identify each RecordComponent are essential to relate successive versions of the data and to link between parts of the EHR;
- attributes that carry clinical semantics (e.g. negation or uncertainty) or parts of the clinical data structure (e.g. the method or rationale for performing a test) are not required in the generic record model but should be part of the SynOD knowledge representation.

More investigation is needed before any attributes can be declared redundant. Clinicians in open fora have requested them all even if they aren't yet employed in practice. As a result future designers of systems should consider a storage structure such as an expandable linked-list of attribute possibilities rather than an explicitly allocated set if space is likely to be at a premium.

#### **CONCLUSION**

The authors have successfully modelled and implemented a full working record server conforming to ENV 13606, capable of storing and retrieving medico-legally complete records, and deployed this in a clinical department. To demonstrate the scalability of the technologies employed, the functioning system was implemented within a busy hospital department to get feedback from clinical staff actually dealing with patient data routinely. It has subsequently been deployed more extensively across north London.

This work has demonstrated that anticoagulant clinic data is inherently hierarchical by extracting the data from the deployed record server and subjecting it to an analysis by RecordComponent type. If it were to be supposed that data was inherently tabular for example, then few data elements would provide contextual information. However in this analysis almost a third of the data existed for no other reason than to establish a hierarchical containment context (i.e. to provide a grouping and heading structure).

It has also been found that much of the richness of the attributes provided were not exploited in this domain. In developing the Synapses architecture, decisions were taken about which features of EHR data were generic or required frequently enough that they should be incorporated within the SynOM, and which were infrequently needed or domain specific or likely to evolve over time and would be best represented as knowledge artefact instances within the Synapses Object Dictionary. The evaluation reported here suggests that the balance might have been struck too far in favour of including attributes within the SynOM. The findings reinforce the need for the main record hierarchy building blocks, but suggest that a simplification of the model attributes might be appropriate.

The process of developing generic EHR information models, arising from research such as the Synapses project and embodied within new-generation interoperability standards and specifications, has been strongly driven by requirements. These requirements have been gathered by the authors over fifteen years, primarily by collecting use cases and examples from clinical communities, and added to successive generations of these models. A priority setting approach has not usually been pursued - all requirements have been received and almost invariably met. This work has shown, admittedly through one example system, how little of the resulting model might actually be needed to represent useful and usable clinical data.

This study has been of value during the drafting of the new CEN/ISO EHR Communications standard (ISO/EN13606 [5]), which retains the principal EHR hierarchy, but possesses fewer contextual attributes as a result.

**SUMMARY**

Known
The need for a structured model to represent federated healthcare data
Candidate models for this, based on a two-tier record-archetype split
Learned
How the models are used in this particular application, reinforcing some design decisions and suggesting changes to others

**CONTRIBUTION OF THE AUTHORS**

Tony Austin wrote the majority of the text and performed the tests described in this paper. Dipak Kalra significantly condensed portions of the text and then added material related to more recent versions of the 13606 standard. Nathan Lea contributed material relating to the numeric data obtained from the operational systems. David Ingram supervised the PhD from which much of the analysis data reported here has been taken. David Patterson is the clinical lead for the anticoagulation service and provided design input to the clinical application and archetypes used by the EHR system.

**REFERENCES**

- [1] Ingram D. The Good European Health Record Project, In: Laires, Ladiera, Christensen Eds, *Health in the New Communications Age*. Amsterdam: IOS Press, 1995; pp. 66-74. ISBN 9051992246.
- [2] Grimson W, Toussaint P, Kalra D, et al. Eds, ODP specification of Synapses. *Synapses Federated Healthcare Record Server* final report. 375 pages. 1999; 66-74. Available from: <https://www.cs.tcd.ie/synapses/public/html/projectdeliverables.html> (last accessed January 2009).
- [3] Kalra D, Austin T, Ingram D, et al. The SynEX project. Current Perspectives in Healthcare Computing '99 Part 1. United Kingdom: BJHC Books 1999; pp. 60-70. ISBN 0-95-35427-0.
- [4] Commission of the European Communities CEN/TC 251, Health Informatics – Electronic Health Record Communication – Part 1: Reference Model. prENV 2003; pp. 13606-1.
- [5] Kalra D, Lloyd D. EN13606 Electronic Health Record Communication Part 1: Reference Model. CEN TC/251, Brussels. February 2007.
- [6] Kalra D, Patterson DLH, Austin T, et al. Synapses in use: supporting cardiac care at the whittington hospital. Proceedings of TEHRE '98. 1998; pp. 306-12.
- [7] Hurlen P, Skifjeld K. Design and functional specification of FHCR server and interfaces. CEC~AIM-synapses federated healthcare Record Server report SPEC 2.1.2. 92 pages. 1997.
- [8] Kalra D, Ed. *Synapses Object Model and Object Dictionary*. CEC~AIM-Synapses Federated Healthcare Record Server report USER 1.3.2. 121 pages. 1997.
- [9] Beale T, Heard S. Archetype Definition Language (ADL) 1.1. openEHR specification. 2004. Available from: <http://www.openehr.org/releases/0.9/architecture/am/adl.pdf> (last accessed January 2009).
- [10] Beale T. Archetypes: constraint-based domain models for future-proof information systems. OOPSLA 2002 Workshop on Behavioural Semantics. Available from: [http://www.deepthought.com.au/it/archetypes/archetypes\\_new.pdf](http://www.deepthought.com.au/it/archetypes/archetypes_new.pdf) (last accessed February 2005).
- [11] Health Level 7 (HL7) Organisation. HL7 Reference Information Model (RIM) 2000.
- [12] Dolin RH, Alschuler L, Behlen F, et al. HL7 Document Patient Record Architecture: An XML Document Architecture Based on a Shared Information Model. Proceedings of the AMIA Annual Symposium 1999; pp. 52-6.
- [13] Chin HL, Wallace P. Embedding guidelines into direct physician order entry: simple methods, powerful results. Proceedings of the AMIA Annual Symposium: 1999; pp. 221-5.
- [14] Kuperman GJ, Teich JM, Tanasijevic MJ, et al. Improving response to critical laboratory results with automation: results of a randomized controlled trial. J Am Med Inform Assoc 1999; 6: 512-22.
- [15] Bates DW, Teich JM, Lee J, et al. The impact of computerized physician order entry on medication error prevention . J Am Med Inform Assoc 1999; 6: 313-21.
- [16] Rind DM, Davis R, Safran C. Designing studies of computer-based alerts and reminders. MD Comput 1995; 12: 122-6.
- [17] Vadher BD. Development and Evaluation of a decision-aid for Warfarin dosing and monitoring during initiation and maintenance therapy. MD Thesis. 1998.
- [18] Vadher B, Patterson DLH, Leaning MS. Evaluation of a decision support system for initiation and control of oral anticoagulation in a randomised trial. BMJ 1997; 314: 1252-6.
- [19] Vadher BD, Patterson DLH, Leaning MS. Validation of an algorithm for oral anticoagulant dosing and appointment scheduling. Clin Lab Haematol 1995; 17: 339-45.
- [20] Vadher BD, Patterson DLH, Leaning MS. Comparison of oral anticoagulant control by a nurse-practitioner using a computer decision-support system with that by clinicians. Clin. Lab. Haem 1997; 19: 203-7.

Received: May 8, 2009

Revised: June 8, 2009

Accepted: June 23, 2009

© Austin et al.; Licensee Bentham Open.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.