

Terminology challenges implementing the HL7 context-aware knowledge retrieval ('Infobutton') standard

Howard R Strasberg,¹ Guilherme Del Fiol,² James J Cimino³

¹Wolters Kluwer Health, Sunnyvale, California, USA

²Department of Biomedical Informatics, University of Utah, Salt Lake City, Utah, USA

³Laboratory for Informatics Development, National Institutes of Health Clinical Center, Bethesda, Maryland, USA

Correspondence to

Dr Howard R Strasberg, 3830 Valley Centre Dr Ste 705, PMB 461, San Diego, CA, 92130, USA; howard.strasberg@wolterskluwer.com

Received 6 August 2012

Accepted 15 September 2012

Published Online First

16 October 2012

ABSTRACT

Point-of-care information needs are common and frequently unmet. One solution to this problem is the use of Infobuttons, which are context-sensitive links from electronic health records (EHR) to knowledge resources, sometimes involving an intermediate broker known as an Infobutton Manager. Health Level Seven (HL7) has developed the Context-Aware Knowledge Retrieval (Infobutton) standard to standardize the integration between EHR systems and knowledge resources. While the standard specifies a set of context attributes and standard terminologies, it leaves to knowledge resources the flexibility to decide how to use these attributes and terminologies to retrieve the most relevant content. This paper describes some of the challenges faced by knowledge resources in trying to locate the most relevant content based on the attribute values for a given Infobutton request. Various approaches to content retrieval are discussed, including the role of indexing with standardized codes, the role of text-based search engines together with their ranking algorithms, and the role of hybrid approaches. Knowledge resource developers must carefully consider business rules, heuristics, and precision/recall tradeoffs when implementing the HL7 Infobutton standard.

INTRODUCTION

Point-of-care information needs are common and frequently unmet.¹⁻⁵ In a 1985 seminal study, Covell *et al* observed that physicians raised two questions for every three patients seen in an outpatient setting.¹ In 70% of the cases these questions were not answered. More recent research has produced similar results,² with little improvement compared with the findings of Covell *et al*. For example, Ely *et al* reported that, in 45% of cases an answer is not pursued at all and, in the remaining 55%, clinicians are still unable to answer 28% of the questions.

Online health knowledge resources are repositories of content containing different kinds of medical information such as from professional to patient, from diagnosis to treatment, from medications to laboratory tests. Examples include PubMed, MedlinePlus, DailyMed, the National Guideline Clearinghouse, as well as various commercial products. Although online health knowledge resources have the answers to most clinicians' information needs,^{6,7} major barriers hinder a more efficient and effective use of these resources.² To overcome these barriers, tools have been designed to help providers quickly identify relevant high-quality knowledge in the context of need.

'Infobuttons' are an example of this kind of tool. Based on the context in an electronic health record (EHR) system (eg, a physician ordering a particular drug for a female patient of child-bearing age), Infobuttons anticipate clinicians' information needs about a given patient and offer links to contextually-relevant knowledge in online resources.⁸ Studies have shown that clinicians enabled with access to Infobuttons were able to meet their information needs in over 85% of the Infobutton sessions, leading to learning or decision enhancement in 62% of these sessions within a median time of 35 s.^{9,10}

Infobutton capabilities are being increasingly supported by knowledge resources and EHR systems¹¹ through standard web services compliant with the Health Level Seven (HL7) Context-Aware Knowledge Retrieval (Infobutton) Standard.¹²⁻¹⁴ The Infobutton standard has been included in the Standards Certification Criteria for the Meaningful Use of EHR Systems recently finalized by the United States Office of the National Coordinator for Health Information Technology.¹⁵ Under these criteria, the use of the HL7 Infobutton standard is required for identifying patient education material and one of the options for identifying diagnostic and therapeutic reference information for linked referential clinical decision support. Note that the criteria require EHR implementers to allow configurable integration with any HL7-compliant knowledge resource, so that meaningful users of health information technology should be able to select the knowledge resources that best meet their needs in each clinical context. In essence, the Infobutton standard provides a standard mechanism for EHR systems to represent and communicate context to knowledge resources. While the standard specifies a set of context attributes and standard terminologies, it leaves to knowledge resource developers the flexibility to decide how to use these attributes and terminologies to retrieve the most relevant content. In fact, interviews with knowledge resource developers who implemented the Infobutton standard consistently revealed the optimal use of standard terminologies to be their main implementation challenge.¹¹ This paper describes some of the challenges faced by knowledge resource providers in the implementation of this standard.

BACKGROUND

Infobutton standard

A typical Infobutton request contains a 'concept of interest' (eg, a laboratory test, medication, or

diagnosis) and some information about the clinical context such as characteristics of the user, the patient, the care setting, and the task being carried out in the EHR. The concept of interest is represented in the HL7 standard as a coded data type class called *mainSearchCriteria*. This class includes a *code*, the code's source *code system* (eg, ICD-9-CM, SNOMED CT), the code's *display name* (textual label) in the source code system, and the *original text* associated with the code and as presented to the EHR user (eg, representing an ICD-9-CM term as 'mainSearchCriteria.v.c=410.00&mainSearchCriteria.v.cs=2.16.840.1.113883.6.103&mainSearchCriteria.v.dn=Acute+Myocardial+Infaction+of+anterolateral+wall,+episode+of+care+unspecified &informationRecipient.languageCode.c=en'). In situations where the EHR does not use a controlled terminology and the concept of interest is a text phrase, HL7 permits code-related parts of the *mainSearchCriteria* value to be blank (eg, sending only the original text).

Examples of contextual parameters include the patient's age group (represented as a MeSH code such as D002648 to represent a child aged 6–12 years) and the user type (eg, nurse, physician). A brief description of the standard is available at http://www.hl7.org/implement/standards/product_brief.cfm?product_id=208. This website also provides a link for HL7 members to download, and for non-members to purchase, the complete standard specification.

Content indexing

Knowledge resource developers have different options for handling coded requests. One option is to index the content using the elements of the Infobutton standard including both standard terminologies and contextual parameters. Another option is to convert incoming query terms (whether text or coded) into search terms (whether text or coded) that are optimized for the resource's existing search engines. The aim of this work is to explore how implementers of the HL7 Infobutton standard might address the various permutations of the mapping of query terms from an Infobutton to search terms used by a resource.

MAPPING INFOBUTTON REQUEST TERMS TO KNOWLEDGE RESOURCE TERMS

Infobutton requests (queries) may contain text terms or standard terminology codes. Text terms may be either narrative text, such as from a problem list, or local controlled terms, such as from a laboratory system or a pharmacy system. Similarly, knowledge resources may be indexed with either text index terms or standard terminology codes. The resulting four permutations are listed in table 1. In this section we will explore each permutation in detail.

Table 1 Permutations for mapping query terms in Infobutton requests to a knowledge resource's indexing terms

Knowledge resource	Infobutton request (query)	
	Text terms	Standard terminology codes (eg, ICD-9-CM)
Text index terms	A	B
Standard terminology index codes (eg, ICD-9-CM)	C	D

Letters A–D refer to subsections of the section on 'Mapping Infobutton request terms to knowledge resource terms'.

A: Text query terms and text search

The simplest mapping process is also the most common situation—text representations of concepts from the EHR being sent to resources that provide only text-based searching functions. While EHRs represent some of their data with controlled terminologies (particularly orderable items like laboratory tests and medications), they often do not use standard or even externally recognizable controlled terminologies. These local terminologies may be represented by codes associated with names or simply by a set of unique names. These names often convey nuances that are specific to the EHR's institution such as 'Unit Dose Ampicillin 500 mg Cap', 'Stat Blood Gluc', 'Smith Pavillion MRI of Chest and/or Abdomen'. As noted above, the HL7 specification allows the use of local terms as *original text* values in the *mainSearchCriteria* parameter.

Resources that provide access to their knowledge through text retrieval, on the other hand, will contain natural language phrases that will be similar to the types of phrases that a human searcher might type into a user interface, but might match poorly to the stylized names used in EHRs. A text phrase can be extracted from the *original text* portion of the *mainSearchCriteria* parameter value, but search engines that seek knowledge that matches all of the words in a search phrase may not be well suited for the local term names from EHRs. Rather, a relaxed matching algorithm that considers partial-word and partial-phrase matches (with results ranked by relevance) may provide a better solution. For example, searching PubMed with the term 'Stat Blood Gluc' (the only text term available from an EHR) returns no results, while searching Google returns 139 000. Translation to a search expression appropriate for PubMed (the text expression 'blood glucose test', which returns 49 267 citations) is not straightforward. It may be easier if the local source provides some mappings to a recognized terminology (such as logical observation identifiers names and codes (LOINC)) that provides better search terms (such as LOINC's related names) but, even then, some of the standard names will be more adequate for searching than others.

B: Standard terminology query codes and text search

With increasing adoption of national standards for clinical data representation, the opportunities for EHRs to insert standard controlled terms and codes into Infobutton links will become easier to accomplish. While there are no regulatory requirements for resource providers to adopt these terminologies, there are likely to be market incentives for resource providers to respond to Infobutton requests in sensible ways. In fact, knowledge resource publishers have been leading the development and implementation of the Infobutton standard as demonstrated in a previous study.¹¹ Resources that continue to support text-only searches will need to use either the *display name* or *original text* portion of the *mainSearchCriteria* class. As with text-to-text mapping described above, relaxed matching algorithms will be needed to provide reasonable performance of the search. However, the choice of *display name* versus *original text* may depend on the terminology in question. For example, a medication represented in RxNorm as 'Amoxicillin trihydrate 600 MG Disintegrating Tablet' (technically this is the *display name*) might be displayed to the user as 'Amoxicillin 660 mg Tab—Disint' (technically this is the *original text*), while a laboratory test represented in LOINC as 'Glucose-SerPl-mCnc' (the *display name*) might be displayed to the users as 'Serum Glucose Test' (the *original text*). In these examples, the resource

might perform best by choosing the *display name* value from RxNorm data and the *original text* value from LOINC data.

Alternatively, resources could translate standard terminology codes into text search terms. For example, a resource might translate ICD-9-CM code 401.1 into ‘benign essential hypertension’. This alternative is explored further in the section below on ‘Search engine ranking: an alternative approach’.

C: Text query terms and standard terminology index codes

When an EHR is only capable of sending text phrases or local terms as values of the *mainSearchCriteria*, interfacing with a resource that requires the use of a standard terminology for retrieval may be problematic. Although some Infobutton Managers provide terminology translation services,^{8 9 16} resource developers cannot assume that this capability will be always available to EHR systems.

A resource that relies solely on obtaining a standard terminology code as input is likely to fail when the original EHR has none to send. If the resource can match on the textual description of the code, there is at least the possibility that the *original text* value from the EHR will match the textual description of a legitimate controlled term. A more robust approach is to use a standard code when available but perform a two-step mapping process when no code is available: first, use string matching to map the EHR term to a standard term and then use the standard term for knowledge retrieval. The MedlinePlus Connect resource from the National Library of Medicine (NLM) uses this approach to map input through its HL7-compliant interface to ICD-9-CM codes. In the non-HL7 world, PubMed has long used a hybrid approach of attempting to match user input behind the scenes to Medical Subject Headings (MeSH) terms where possible.

D: Standard terminology query codes and standard terminology index codes

The Holy Grail of semantic interoperability is the ability for the sender and receiver of information to represent data with standard terminology such that the receiver understands the sender’s intended meaning. Infobutton implementations are becoming available where the sender and receiver are using the same controlled standard terminology, perhaps facilitated by a growing adoption of standard terminologies in general by EHR systems. However, unless there is an exact match between the coded term from the EHR and the coded term used to index the knowledge in the resource, some additional mapping or a set of heuristics using standard terminologies will be needed.

The next section of this paper explores some of the complexities of this process and some of the potential solutions.

The problem is compounded when the sender and receiver use different standard terminologies. Some success with automated translation between standard terminologies has been achieved by various researchers, especially through the use of resources such as the NLM’s Unified Medical Language System (UMLS)¹⁷ and the Open Biomedical Ontologies.¹⁸ Indeed, one of the earliest implementations of Infobuttons used the UMLS to map between ICD-9-CM codes in an EHR and MeSH terms in a Medline search engine.¹⁹

The ultimate solution to this problem may be the inclusion of generic (any-to-any) terminology translation services within the Infobutton Manager rather than expecting resources to provide this function. However, given the large number of standard terminologies in use, a truly generic (any-to-any) solution would be quite complicated to develop, although it could be facilitated by the UMLS. Until then, Infobutton implementers will consider the terminologies available in their EHRs, the terminologies recognized by available knowledge resources, and provide links accordingly. The task of using the standard controlled terms for retrieval will, however, always fall to the knowledge resource, as described in the next section.

ADVANCED APPROACHES TO CONTROLLED TERM MAPPING

Challenges with indexing diagnoses

Suppose that a knowledge resource provider carefully codes its content with the most appropriate ICD-9-CM codes, perhaps even using human expert coders. However, unless the coders create an index with all existing ICD-9-CM codes, handling ICD-9-CM requests that do not exactly match the ICD-9-CM codes assigned to the content will remain an issue. Table 2 shows some illustrative examples of different types of inexact matches.

As shown in table 2, implementers must make decisions on how their resources should behave in each of the circumstances represented by the examples. When is an inexact match better than no match? When is it worse? When might it be harmful? In the example involving codes 401.9 (unspecified essential hypertension) and 401.1 (benign essential hypertension), matching on the sibling code may be reasonable. In the example involving codes 250.01 (type 1 diabetes mellitus) and 250.02 (type 2 diabetes mellitus), however, such a match is probably not appropriate. The creation of matching rules that work for the general case therefore remains challenging.

Table 2 ICD-9-CM matching issues with examples

Challenge	Infobutton request code	Document index code	Comments
Request code is a descendent of a document code	493.21 (chronic obstructive asthma with status asthmaticus)	493.2 (chronic obstructive asthma)	Match the parent?
Request code is an ancestor of a document code	493.2 (chronic obstructive asthma)	493.21 (chronic obstructive asthma with status asthmaticus)	Match the child?
Request code is a sibling of a document code, but the two codes are not mutually exclusive	401.9 (unspecified essential hypertension)	401.1 (benign essential hypertension)	Match the sibling?
Request code is a mutually exclusive sibling of a document code	250.01 (type 1 diabetes mellitus)	250.02 (type 2 diabetes mellitus)	Not appropriate to match the sibling
Request code is a descendent of a code in one document and an ancestor of a code in another document	493.2 (chronic obstructive asthma)	Document 1: 493.21 (chronic obstructive asthma with status asthmaticus) Document 2: 493 (asthma)	If the desire is to go directly to a document, would it be better to match the child, match the parent, match both, or stop to ask the user?

Numbers represent hierarchical ICD-9-CM codes, with ancestor codes having descendant codes with additional rightward digits.

Creating matching rules for specific cases could theoretically be done, but the effort involved would be significant. For example, a rule that tried to locate the most specific document based on the requested code (but nothing more specific than the requested code) would not address sibling matches and would exclude child matches, even though some child matches may be appropriate.

In cases where more than one close but inexact match is available, users may be given a list of choices. However, the choices must still be ranked in some order, so rules are needed for prioritizing different ‘close’ matches.

Challenges with indexing medications

RxNorm was created by the NLM as a standardized nomenclature for drugs. The idea sounds simple enough, but in practice it is fairly complex. RxNorm contains several different term types, such as *Ingredient* (eg, 1202; Atenolol (IN)), *Semantic Clinical Drug* (eg, 197379; Atenolol 100 MG Oral Tablet (SCD)), *Semantic Branded Drug* (eg, 104305; Atenolol 100 MG Oral Tablet (Totamol) (SBD)), *Semantic Clinical Drug Form* (eg, 370619; Atenolol Oral Tablet (SCDF)) and *Semantic Clinical Drug Component* (eg, 315436; Atenolol 100 MG (SCDC)). In fact, RxNorm contains 59 codes for the drug atenolol, excluding drug combinations.

To provide complete support for RxNorm, a resource provider with knowledge on atenolol would have to accept all 59 codes, even though RxNorm was designed to simplify drug nomenclature. There are several ways a knowledge resource could accommodate all these different codes. One approach would be to index the content manually with all applicable RxNorm codes. A second approach would be to index the content with a single representative RxNorm code, but to precompile a list of all related RxNorm codes when building the index for the retrieval system. A third approach would be to operate a terminology server that could map between different types of RxNorm codes at run time. The NLM’s RxNorm Application Program Interface may be useful at run time for this purpose.

Challenges with indexing laboratory test results

LOINC suffers from the same problem as RxNorm in that multiple LOINC codes represent a single laboratory test. For example, there are six different LOINC codes for a serum potassium test but, from a clinical standpoint, the distinction among these codes is often not relevant. In particular, there are codes for results as a substance concentration (2823-3), a mass concentration (22760-3), either a substance or a mass concentration (42569-4), a second specimen (12812-4), a third specimen (12813-2), and post-dialysis (29349-8).

Once again, resource providers would have to code their content using multiple codes or else develop some internal equivalency mapping database. We are unaware of an existing database that would, for example, assert that LOINC codes 12812-4 and 2823-3 are equivalent for knowledge retrieval purposes. In this respect, LOINC is different from RxNorm since RxNorm contains tables that explicitly enumerate the relationships between different RxNorm codes.

Ranking coded search results

In the simple case of a request for code X and multiple documents in the corpus indexed with code X, the knowledge resource developer has to find a way to rank the results. This question will be explored further in the sections below on ‘Search engine ranking: an alternative approach’ and ‘Hybrid approach’.

In more complex cases, requests may contain a variety of contextual information. If a knowledge resource has documents that match the requested main search criterion and all of the contextual information, then clearly it makes sense to display these documents to the user. On the other hand, very often only some of the contextual information can be matched. The question then becomes one of how to prioritize the various contextual attributes and whether the rules for prioritization vary depending on the circumstances.

For example, consider a request from a nurse for knowledge on ICD-9-CM code 493.2 (Chronic Obstructive Asthma). Suppose the following:

- ▶ No document is available that is both written for a nurse and is coded with ICD-9-CM code 493.2.
- ▶ A document (Document 1) is available that is written for a nurse and is coded with ICD-9-CM code 493.
- ▶ Another document (Document 2) is available that is written for a physician and is coded with ICD-9-CM code 493.2.

Which document would users prefer in this situation?

As another example, consider a request from a physician for knowledge on ICD-9-CM code 493.2 where the patient is a child. Suppose the following:

- ▶ No document is available that contains knowledge for ICD-9-CM code 493.2 where the patient is a child.
- ▶ A document (Document 3) is available that contains knowledge for ICD-9-CM code 493 where the patient is a child.
- ▶ A document (Document 4) is available that contains knowledge for ICD-9-CM code 493.2 where the patient is an adult.

Once again, which document would users prefer?

Finally, consider the combination of these scenarios—a nurse requests knowledge for ICD-9-CM code 493.2 where the patient is a child. How should the documents in table 3, none of which exactly matches the request, be ranked?

The challenges illustrated in this section may be addressed by developing a ranking algorithm. For example, inexact matches may get less weight than exact matches and weighted scores could be assigned to various contextual parameters. Knowledge resource developers would have to decide how to weight the different components of such an algorithm. For example, they would have to decide on the relative penalty in the final ranking score between a provider mismatch, an age mismatch, and an inexact code match.

Search engine ranking: an alternative approach

Given the challenges involved in coding content (coded indexing approach) with standard terminologies and handling requests by finding documents that match (on either an exact or a fuzzy basis) the requested code, resources may consider a different approach which is to use their existing search engine

Table 3 Partial matches to a request containing ICD-9-CM 493.2, nurse, and child

Document	Provider	Age	ICD-9-CM Code
5	Nurse	Adult	493.2
6	Nurse	Adult	493
7	Nurse	Child	493
8	Physician	Adult	493.2
9	Physician	Adult	493
10	Physician	Child	493.2
11	Physician	Child	493

(resource's search engine approach). In this scenario, the knowledge resource would maintain a mapping file to map incoming standardized codes into search queries optimized for its search engine. In effect, the knowledge resource developer would be making a deliberate decision to adopt the model in the subsection 'Standard terminology query codes and text search' instead of the model in the subsection 'Standard terminology query codes and standard terminology index codes', even if the latter were otherwise feasible.

For example, in a coded indexing approach, five documents may be indexed with ICD-9-CM code 401.1 (benign essential hypertension). As noted above, it is unclear how to rank these results using an indexing approach. However, if the code was converted to a search query such as benign essential hypertension, the knowledge resource could leverage the existing ranking capabilities of its search engine in order to return results in a ranked order. Such capabilities might include the use of term frequency and inverse document frequency, the use of some variant of link analysis, and the use of log analysis. The details of these search engine ranking methods are beyond the scope of this article, but interested readers can consult the references by Salton *et al*, Page *et al*, and Joachims *et al*, respectively, for additional information.²⁰⁻²²

The use of the resource's search engine in lieu of the coded elements of an Infobutton request can impact precision and recall. For example, consider a document primarily about hypertension and coded with ICD-9-CM 401.1. This document may mention sleep apnea (ICD-9-CM code 327.23) as a secondary cause of hypertension but it is not indexed with this code (327.23). In a coded indexing approach, an incoming request for 327.23 would therefore return no results. In the resource's search engine approach, however, code 327.23 may be mapped to the query term 'sleep apnea' which the search engine would find in this document and therefore return this document. Therefore, compared with the coded indexing approach, the resource's search engine approach may increase recall and decrease precision.

Hybrid approach

We have thus far considered both coding and searching approaches independently. Consider two documents, each coded with ICD-9-CM 431 (intracerebral hemorrhage), but with one really about cortical intracerebral hemorrhage and the other really about intracerebral hemorrhage in the cerebellum. An incoming request for ICD-9-CM code 431 would be matched to both documents (using the model in the subsection 'Standard terminology query codes and standard terminology index codes'). Suppose, however, that the request also contains the original text (as entered into the EHR) 'intracerebral hemorrhage in cerebellum'. Use of the original text by a search engine would identify the correct document (using the model in the subsection 'Text query terms and text search'). Since the original text can be more granular than the code, its use should be strongly considered. Parallel use of the code is also recommended to avoid text-based matches to otherwise irrelevant documents.

Note that while combining the code and the original text appears to make sense, combining the code and the display name may be less useful. The display name is simply a string expression of the code whereas the original text is the actual text from the EHR and may have different granularity. On the other hand, the original text may contain institution-specific information, non-standard abbreviations, or even misspelled terms that may not be adequate search terms. As noted in the subsection 'Text query terms and standard terminology index

codes' above, whether the display name or original text is preferred may depend on the standard terminology being used. Knowledge resource developers should therefore consider business rules governing when to use the original text, the display name, or both. The display name should obviously be used if no other information is provided.

DISCUSSION

This paper explores some of the challenges faced by knowledge resource developers when implementing the Context-Aware Knowledge Retrieval (Infobutton) standard. Implementing this standard is not as simple as indexing some content with a few standardized codes and assuming that retrieving direct matches to these codes will produce an optimal set of results. Instead, a variety of subtleties must be considered, such as how to handle inexact matches, how to handle different term types, and how to rank the results. A variety of approaches can be used, including retrieving exact code matches, retrieving related code matches and, through query expansion, leveraging the ranking capabilities of search engines. We do not recommend a single approach; instead, we recommend that implementers consider the issues raised in this paper and decide on the best approach for each specific implementation.

The forthcoming requirements to replace ICD-9-CM with a combination of SNOMED CT and ICD-10-CM will likely be a double-edged sword for knowledge resource developers. On one side of the coin, since these newer terminologies are much more granular than ICD-9-CM, whenever there are exact matches between coded Infobutton requests and similarly coded clinical decision support content, users are likely to benefit from more relevant and more focused information. On the other hand, given the sheer magnitude of the terms and concepts in these newer terminologies, exact matches will probably occur less frequently, thereby exacerbating the issues described earlier in this paper concerning how to handle inexact matches.

To address some of the challenges described herein, there may be a role for standardized value sets such as those found in the Centers for Medicare and Medicaid Services (CMS) quality measures (www.cms.gov) and those to be available from the NLM Value Set Authority Center.²³ Value sets would define a set of codes for clinical concepts of interest such as hypertension, β blockers, or hypokalemia. These value sets could be used in content indexing to improve the retrieval results. However, the value sets described above are being created to support quality measures, and it is unknown how useful these value sets will be for information retrieval purposes. In addition, for large terminologies, the use of subsets such as SNOMED CT Core by both parties would improve the likelihood of exact matches between EHRs and knowledge resource content.

This work is limited by the experience of its authors; however, collectively they have over 40 years of experience spanning numerous Infobutton and Infobutton Manager implementations. They have also been instrumental in the development of the HL7 Infobutton standard.

Optimal use of the HL7 Infobutton standard, including use of its numerous contextual parameters, would be expected to improve precision compared with traditional searching without being able to specify the context. Future work will need to study this question further and, in particular, it will need to involve formal precision/recall analysis of different indexing and retrieval strategies. Until that work is completed, individuals responsible for Infobutton implementations at provider organizations should consider precision and recall tradeoffs when selecting Infobutton resources for various points in the

user workflow. In addition, future work should compare generic strategies for all types of queries with approaches that may differ based on specific topic areas and/or use cases. There may also be a role for professional societies to assist with the development of optimal indexing and retrieval strategies for content in their respective areas of expertise.

CONCLUSION

Knowledge resource developers implementing the HL7 Infobutton standard must deal with multiple challenges, particularly related to differences between requested codes and codes assigned as content index terms. Future research is needed to develop heuristics around the handling of close but inexact code matches. Attention needs to be paid to parent, child, and sibling relationships, as well as to the development of value sets of codes that can be considered equivalent for knowledge retrieval purposes.

Contributors All three authors made substantial contributions to the paper. HRS conceived of the idea and drafted the 'Background', 'Advanced approaches to controlled term mapping', 'Discussion' and 'Conclusion' sections. GDF drafted the 'Introduction' section. JJC suggested and drafted the 'Mapping Infobutton request terms to knowledge resource terms' section. All authors reviewed all sections of the paper.

Funding This project was supported in part by grant number K01HS018352 from the Agency for Healthcare Research and Quality. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Agency for Healthcare Research and Quality. This research was also supported in part by intramural research funds from the NIH Clinical Center and the National Library of Medicine.

Competing interests None.

Provenance and peer reviewed Not commissioned; externally peer reviewed.

REFERENCES

- Covell DG, Uman GC, Manning PR. Information needs in office practice: are they being met? *Ann Intern Med* 1985;**103**:596–9.
- Ely JW, Osheroff JA, Chambliss ML, et al. Answering physicians' clinical questions: obstacles and potential solutions. *J Am Med Inform Assoc* 2005;**12**:217–24.
- Gorman P. Information needs in primary care: a survey of rural and nonrural primary care physicians. *Stud Health Technol Inform* 2001;**84**(Pt 1):338–42.
- Ramos K, Linscheid R, Schafer S. Real-time information-seeking behavior of residency physicians. *Fam Med* 2003;**35**:257–60.
- Smith R. What clinical information do doctors need? *BMJ* 1996;**313**:1062–8.
- Magrabi F, Coiera EW, Westbrook JI, et al. General practitioners' use of online evidence during consultations. *Int J Med Inform* 2005;**74**:1–12.
- Westbrook JI, Coiera EW, Gosling AS. Do online information retrieval systems help experienced clinicians answer clinical questions? *J Am Med Inform Assoc* 2005;**12**:315–21.
- Cimino JJ, Elhanan G, Zeng Q. Supporting Infobuttons with terminological knowledge. Proceedings: Conference of the American Medical Informatics Association/AMIA Annual Fall Symposium AMIA Fall Symposium. 1997:528–32.
- Del Fiol G, Haug PJ, Cimino JJ, et al. Effectiveness of topic-specific Infobuttons: a randomized controlled trial. *J Am Med Inform Assoc* 2008;**15**:752–9.
- Maviglia SM, Yoon CS, Bates DW, et al. KnowledgeLink: impact of context-sensitive information retrieval on clinicians' information needs. *J Am Med Inform Assoc* 2006;**13**:67–73.
- Del Fiol G, Huser V, Strasberg HR, et al. Implementations of the HL7 Context-Aware Knowledge Retrieval ('Infobutton') Standard: challenges, strengths, limitations, and uptake. *J Biomed Inform* 2012;**45**:726–35.
- Del Fiol G. Context-aware Knowledge Retrieval ('Infobutton'), Knowledge Request Standard. Health Level 7 International, 2010. http://www.hl7.org/v3ballot2010may/html/domains/uvds/uvds_Context-awareKnowledgeRetrieval%28Infobutton%29.htm (accessed 16 Jan 2012).
- Del Fiol G. Context-aware Knowledge Retrieval ('Infobutton'), URL-Based Implementation Guide Standard. Health Level 7 International, 2010. http://www.hl7.org/v3ballot2010may/html/domains/uvds/Infobutton_IG.pdf (accessed 16 Jan 2012).
- Del Fiol G. Context-aware Knowledge Retrieval ('Infobutton'), Service-Oriented Architecture Implementation Guide Draft Standard. Health Level 7 International 2011. http://www.hl7.org/v3ballot/html/domains/uvds/V3_IG_SOA_KM_INFobutton_R1_D2_2011JAN.zip (accessed 16 Jan 2012).
- Department of Health and Human Services. Health Information Technology: Standards, Implementation Specifications, and Certification Criteria for Electronic Health Record Technology, 2014 Edition; Revisions to the Permanent Certification Program for Health Information Technology, 2012. <http://www.gpo.gov/fdsys/pkg/FR-2012-03-07/pdf/2012-4430.pdf> (accessed 2 Jul 2012).
- Del Fiol G, Kawamoto K, Cimino JJ. Open-source, standards-based software to enable decision support. *AMIA Annu Symp Proc* 2011;2127.
- Patel CO, Cimino JJ. A network-theoretic approach for decompositional translation across open biological ontologies. *J Biomed Inform* 2010;**43**:608–12.
- Kawamoto K, Lobach DF. Design, implementation, use, and preliminary evaluation of an UMLS-enabled terminology web service for clinical decision support. *AMIA Annu Symp Proc* 2006:979.
- Cimino JJ, Johnson SB, Aguirre A, et al. The Medline Button. In: Frisse ME, ed. Proceedings of the Sixteenth Annual Symposium on Computer Applications in Medical Care. Baltimore, MD, November. New York: McGraw-Hill, 1992:81–5.
- Salton G, Buckley C. Term-weighting approaches in automatic text retrieval. *Info Proc Mgmt* 1988;**24**:513–23.
- Page L, Brin S, Motwani R, et al. The PageRank Citation Ranking: Bringing Order to the Web. Technical Report. Stanford InfLab. 1999. <http://ilpubs.stanford.edu:8090/422/1/1999-66.pdf> (accessed 5 Sept 2012).
- Joachims T, Granka L, Pan B, et al. Accurately determining clickthrough data as implicit feedback. Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. 2005. <http://www.nlm.nih.gov/healthit/vsac.html> (accessed 8 Oct 2012).