CONFERENCE: The Ninth International Symposium for Health Information Management Research (iSHIMR2004)

TITLE: Augmenting GEM-Encoded Clinical Practice Guidelines with Relevant Best-Evidence Autonomously Retrieved from MEDLINE

AUTHORS: Michael Kershaw, Syed Sibte Raza Abidi, Evangelos Milios

KEYWORDS: Clinical Practice Guidelines, Best Evidence, Literature Search, Search Strategy, MEDLINE

PAPER TYPE: Long Paper

CORRESPONDING AUTHOR DETAILS

NAME: Dr. Syed Sibte Raza Abidi

ADDRESS: Faculty of Computer Science, Dalhousie University, Halifax B3H 1W5, Nova Scotia, Canada

EMAIL: sraza@cs.dal.ca

PHONE: 1-902-4942129

FAX: 1-902-4921517

AUGMENTING GEM-ENCODED CLINICAL PRACTICE GUIDELINES WITH RELEVANT BEST-EVIDENCE AUTONOMOUSLY RETRIEVED FROM MEDLINE

Michael Kershaw, Syed Sibte Raza Abidi, Evangelos Milios

Health Informatics Laboratory, Faculty of Computer Science, Dalhousie University, Halifax, Canada.

Abstract

Clinical practice guidelines (CPG) are instrumental in standardizing the quality and delivery of care across different practitioners, departments and even institutions. Health practitioners when working with CPG like to reflect on the current best evidence either to validate or to supplement their understanding of the CPG. In this paper we investigate the potential of supplementing computerized CPG with current and relevant best-evidence sourced from reliable medical literature repositories. We present a web-enabled Best-evidence Retrieval and Delivery (BiRD) system that provides the functionality to autonomously retrieve pertinent medical literature with respect to user-specified content from a GEM-encoded CPG. We have developed a multi-level literature search strategy that both categorizes the search query towards a priori defined clinical query intentions, and subsequently filters insignificant medical terms from the search query. The resultant is a highly focused medical literature search query that is objectively derived from CPG content. The technical architecture comprises existing medical language processing tools and vocabularies, together with newly developed tools to (a) generate optimum search queries; (b) retrieve medical articles from MEDLINE; and (c) embed the retrieved medical articles within XML-based CPG.

KEYWORDS: Clinical Practice Guidelines, Best Evidence, Information Retrieval, Search Strategy, MEDLINE

1. INTRODUCTION

Evidence-based healthcare is a prevalent practice amongst both medical practitioners and management as it provides a sound basis for quality and consistent healthcare delivery. In this regard, Clinical Practice Guidelines (CPG)—defined as "systematically developed statements to assist practitioners and patient decisions about appropriate health care for specific circumstances" [1]—provide means to improve the quality, effectiveness, and standardization of healthcare practices across different practitioners, departments and even institutions. CPG describe the most appropriate diagnostic and treatment plan based on both the best scientific evidence and expert consensus available at the time of its compilation [2].

Lately, there has been an increase in initiatives to computerize CPG so that they can be incorporated in clinical and diagnostic-support systems and be used at the point of care [3, 4]. From an evidence-based healthcare perspective, Computerized CPG (C-CPG) are able to provide rapid evidence-based recommendations about what medical procedures to perform based on an individual patient's health profile [5,6,7]. A number of C-CPG representation formalisms have been outlined by researchers and successfully deployed in medical systems.

Notwithstanding the validity and importance of CPG, yet medical practitioners when working with C-CPG have a tendency to reflect on the current best evidence primarily to validate or to supplement their understanding of the CPG [8]. The idea is to refer to current medical literature to acquire useful insights into the rationale, past clinical trials, the diversity of opinions, related observations and current clinical evidence with respect to certain recommendations/procedures/treatments specified in the CPG. To achieve the above, typically health practitioners attempt to retrieve medical literature from on-line medical repositories, such as MEDLINE, Cochrane, etc. Retrieving the 'right' and 'relevant' medical literature is a subjective process, largely contingent on how best the practitioner frames both the medical problem and the clinical intention within the search query. Typically, the user may provide a list of medical terms, as the basis for the literature search, as per his/her understanding of the medical problem and ability to specify the clinical intention.

We argue that since C-CPG are highly focused in their treatment of the problem, it is important that the best evidence reviewed by practitioners to supplement their understanding of the C-CPG should also be equally focused towards either the entire C-CPG or a particular segment of the C-CPG. We contend that the quality of best evidence provided to practitioners can be improved if the origin of the literature search query is not subjective, rather it should be objective in the sense that it is derived from the original text of the CPG segment. This is an interesting and practical literature search alternative. Since C-CPG contain elements of best evidence in a semantically-explicit and syntactically-formal format, it is therefore opportunistic to use the original C-CPG content as the basis for searching and retrieving medical literature that is more focused and relevant towards the respective C-CPG content. We believe that C-CPG content can be used to derive an objective search query to realize optimal medical literature search.

Given the need for timely best evidence in concert with a computerized CPG, in this paper we investigate the potential for optimal medical literature search based on C-CPG. From the perspective of the health practitioner, we present a web-enabled *Best-evidence Retrieval and Delivery* (BiRD) system that provides the functionality to autonomously retrieve pertinent medical literature with respect to user-specified content (this may be the title, statements, paragraph, section) from a GEM-encoded C-CPG [9]. The idea is to supplement computerized CPG with current and relevant best-evidence sourced from reliable medical literature repositories. Note that, GEM (Guideline Elements Model) is an XML-based guideline document model that consists of more than 100 discrete tags to represent the many aspects of a CPG.

We present a multi-level literature search strategy that both categorizes the search query towards a priori defined clinical query intentions, and also filters insignificant (in terms of searching the medical literature in the given context) medical terms from the search query. The resultant is a highly focused medical literature search query that is objectively classified to a query type and comprises a relatively short list of significant medical terms. The technical architecture comprises the application of existing medical language processing tools and vocabularies, together with newly developed tools to generate optimum search queries, to retrieve medical articles from MEDLINE and to systematically deliver the retrieved medical articles within existing XML-based CPG.

2. SOLUTION DESIGN

Designing a functional solution to supplement C-CPGs with relevant medical literature involved the following tasks: (1) Allowing users to identify the C-CPG content for which corresponding medical literature is sought; (2) Determining the medical terms, from the selected C-CPG content, to be used to generate a search query for literature retrieval; (3) Converting the identified medical terms to MeSH compliant terms because medical literature at MEDLINE is indexed based on MeSH terms; (4) Generating a context-sensitive medical literature search query (comprising a set of MeSH terms) to retrieve the most relevant medical articles; (5) Presenting the search query to MEDLINE in compliance with the operational limits enforced by PubMed, to retrieve the most relevant medical articles; (6) Analyzing the search query results to ensure optimum coverage and relevance of the query results; and finally (7) Returning the retrieved medical articles to the corresponding segments within the C-CPG.

2.1. Search Query Generation Strategy

For maximum impact it is important that the evidence provided to practitioners is relevant—i.e. the medical articles retrieved from MEDLINE are in line with both the intention and context of the C-CPG segment in question.

The premise of our search strategy is that:

(a) The search query should represent the most likely underlying intention/theme of the selected medical content. One way is to categorize the query to a query type based on a set of a priori defined query types [10]. The query type should be objectively determined, as opposed to be subjectively assigned by a user, based on the nature and meaning of candidate MeSH terms derived from the medical content;

(b) The search query should be a combination of the query type and a list of candidate MeSH terms. For instance, if the query type is determined to be *etiology* and the candidate terms are *{Kidney, Urinary Tract Infections, Radiation}* then the medical articles retrieved should have a focus on the etiology of as opposed to the diagnosis or treatment of the same;

(c) The search query should only include a small number of highly significant MeSH terms. A large number of terms tend to make the query too specific which usually results in no medical articles being retrieved. This therefore calls for the need to filter out less significant MeSH terms from the set of candidate MeSH terms.

Our medical literature search strategy attempts to generate an optimum search query by (a) using MeSH terms only [11]; (b) categorizing the search query to a query type for a more focused search [10]; and (c) reducing the length of the search query by removing insignificant medical terms. Categorization of the search query significantly reduces the search space, thus allowing the search engine to focus on a smaller set of medical articles that are indexed under the search category in question. Reducing the length of the search query has two implications; (i) the search query becomes more generic, yet more focused, (ii) the search query is less constrained as it does not need to satisfy a larger list of search terms, where some of the search terms were not significant yet impacting the granularity of the search query.

In our work, we have devised a multi-level query generation model that comprises two distinct stages: (1) *Query categorization stage* and (2) *Term filtering stage*.

2.1.1. Query Categorization Stage—Step 1: We begin by categorizing the query into four query types:

- *Therapy*: Content to be retrieved relates to therapy, prevention or rehabilitation of a disease or condition.
- *Diagnosis*: Content to be retrieved relates to evaluation of a disease or condition.
- *Etiology*: Content to be retrieved relates to causation of a disease or condition.
- *Prognosis*: Content to be retrieved predicts or forecasts the course of a disease or condition.

The above query classification scheme has been suggested as PubMed's Research Methodology Filters [10]. Each query type is associated with a set of UMLS semantic types; medical terms belonging to these UMLS semantic types are regarded as *trigger words*—i.e. medical terms that intrinsically reflect the intention of the query type. For instance, the medical terms 'catheterization' belongs to the UMLS semantic type 'therapeutic or preventive procedure' which is related to the query type 'therapy'. Hence, the presence of the medical term 'catheterization' in the selected C-CPG content strengthens the belief that the likely focus of the C-CPG content is 'therapy', and the query type should be set as *therapy*. The set of semantic types representative of each query type have been determined by medical researchers via surveying clinicians, medical librarians and researchers in Canada and the United States [10].

Given a set of candidate medical terms, we determine the most appropriate query type by comparing the candidate medical terms against the set of trigger words for each query type. The query type associated with the most matched trigger words is deemed as the target query type for the candidate medical terms.

2.1.2. Query Categorization Stage—Step 2: The second level of categorization is based on the age dimension. Medical studies are specific to age and the results of such studies are applicable to a definitive age-group, hence it is important to retrieve and provide evidence that is pertinent to the age-group in question.

For our purposes, we categorize the query into four age groups: *infant* (0 - 23 Months), *child* (0 – 18 years), *adult* (19+ years), *elderly* (65+ years). The age-groups were derived by subsuming 13 different age-groups defined by PubMed to classify queries. The trigger words for each age group were derived using the UMLS semantic types related to the age dimension—each MeSH term belonging to the Age Group UMLS semantic type was associated with an age-group and it served as the trigger word that age-group. For instance, the MeSH terms (in the Age Group UMLS semantic type) {Infant; Infant, low birth weight; Infant, newborn; Infant (postmature); Infant (premature); Infant (small for gestational age)} would serve as the trigger word for the age-group *infant*. Whereas, the MeSH terms {Elderly; Centenarians; Nonagenarians; Octogenarians; Frail elderly} would trigger the age-group *elderly*.

2.1.3. Term Filtering Stage—Step 1: Categorization of the search query leads to the next stage in which we need to identify and then subsequently filter-out those medical terms that are potentially insignificant and thus unnecessarily constrain the query. We have designed three (3) term filters which are used in sequence; each filter is used only if the number of candidate query terms is greater than a user-defined *query length threshold*.

The first term-filter is designed to remove all candidate query terms that belong to a list of filterable UMLS semantic types. These UMLS semantic types are deemed to be of less medical significance [11], for instance the terms listed under the semantic type Reptile are seemingly of nominal significance.

The default list of filterable semantic types is {Physical Object, Amphibian, Plant, Alga, Animal, Vertebrate, Invertebrate, Fish, Bird, Reptile, Research Device, Idea or Concept, Temporal Concept, Qualitative Concept, Quantitative Concept, Functional Concept, Spatial Concept, Geographic Area, Intellectual Product, Classification, Regulation or Law, Language, Occupation or Discipline, Organization, Professional Society, Self Help or Relief Organization, Group, Population Group, Family Group, Professional or Occupational Group}.

In practice, users are allowed to modify the list to retain the semantic types that they deem as being significant for the problem at hand. Furthermore, two considerations are taken into account before filtering the candidate query term: (1) If the frequency of the term belonging to one of the filterable semantic type is greater than a pre-defined threshold we regard the term to have some significance in the context of the medical content in question. In this case the term is not filtered; (2) If a candidate query term belongs to multiple semantic types, and only one of the semantic types is a filterable semantic type then the term is not filtered.

2.1.4. Term Filtering Stage—Step 2: The second term-filter is an extreme case of the first term-filter. It removes a candidate query term if any one of its semantic types belongs to the filterable semantic type group.

2.1.5. Term Filtering Stage—Step 3: We know that the candidate query terms are derived from the CPG content in question. In practice, a text parsing process identifies the noun phrases in the CPG content, and subsequently translates the words within the noun phrase to corresponding MeSH terms. In this process, each noun phrase may map to multiple MeSH terms; each MeSH term is assigned a *mapping score* (out of 1000) that indicates the strength/confidence of the mapping of the original term to a corresponding MeSH term. At times, a single noun phrase may map to multiple MeSH terms, each MeSH term having a different mapping score. For instance the noun phrase "kidney cortex" is mapped to the following three MeSH terms, each with a different mapping score (shown in the square brackets): [1000] Kidney Cortex; [755] Cascara; and [694] Kidney.

If a noun phrase produces multiple MeSH terms, in our scheme we are interested in retaining only the highest mapping score MeSH term. The third term-filter is designed to filter-out all terms with a low mapping score hat originate from the same noun-phrase. So, for the above example, the third term-filter will only retain the MeSH term "kidney cortex" and filter-out the terms "cascara" and "kidney".

3. METHODS

Our BiRD system features the application of existing medical language processing tools, medical vocabularies and utilities to access PubMed. In addition, we have developed modules to (i) input and visualize a GEM-encoded C-CPG (or text documents), (ii) specify C-CPG content for which evidence is sought, (iii) autonomously generate an optimal search query as per the above-mentioned search methodology, (iv) present the query to PubMed in accordance with PubMed's querying regulations, (v) retrieve the medical articles, and (vi) incorporate the medical articles within the C-CPG (next to the selected C-CPG content). Figure 1 shows the functional diagram and the constituent modules of BiRD.

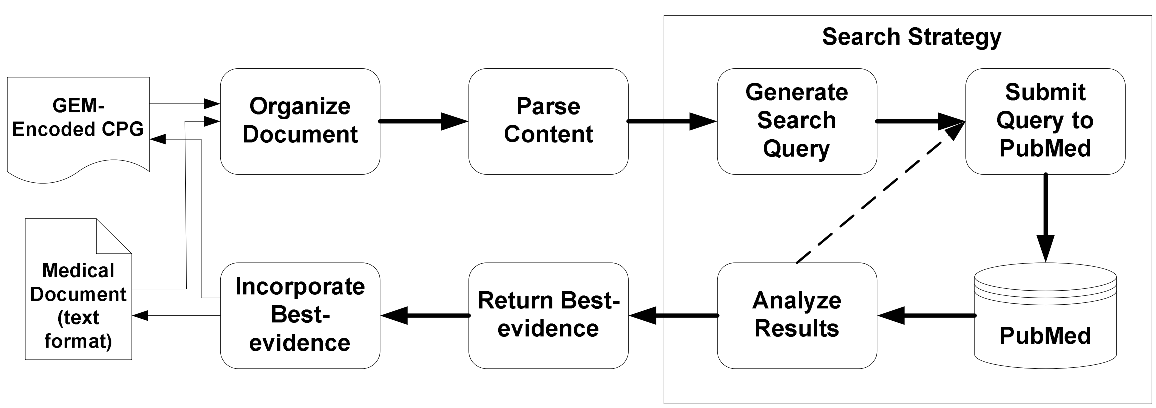

Figure 1: The functional overview of BiRD.

In the forthcoming discussion we will discuss the functionality of the various components of the BiRD system using a working example of a GEM-encoded CPG for managing urinary tract infections.

3.1. C-CPG Visualizer

The functionality of the C-CPG visualizer is to present a GEM-encoded C-CPG to the user to specify the C-CPG content of interest. A GEM document is in XML format, comprising a number of semantically-significant tags that help organize the entire CPG content into meaningful blocks. The C-CPG visualizer leverages the GEM tags to parse the C-CPG into blocks of C-CPG content, each demarcated by a set of GEM tags.

Users can not only view the C-CPG but also simultaneously select those C-CPG content blocks for which supplementary best evidence is required. As shown in figure 2, users can select all or specific C-CPG blocks by clicking the 'check box' next to the C-CPG content of interest.

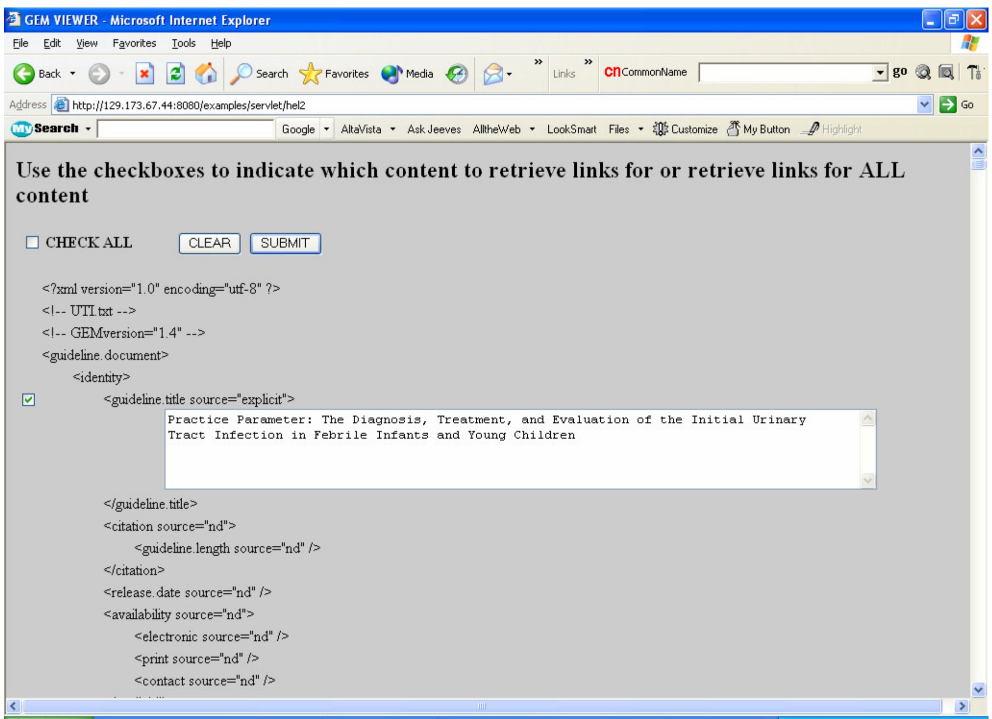

Figure 2: Screen shot of the user-interface to specify the search query input. The user is asked to select the CPG content for which he/she wants to retrieve best evidence.

3.2. MeSH Term Generator

This module is responsible for generating MeSH terms from the original text of the selected C-CPG content. We use the MetaMap Transfer (MMTx) tool, developed by the National Library of Medicine, that offers the functionality to translate plain text to meaningful MeSH terms. We will not be able to provide details of the working of the MMTx tool, however in figure 3 we show a C-CPG content block and the resulting MeSH terms returned by MMTx. The C-CPG content is taken from a urinary track infection CPG available at http://ycmi.med.yale.edu/GEM.

We calculate the frequency of each MeSH term identified by MMTx in order to establish the importance of the term within the given C-CPG content. Table 1 shows the MeSH terms derived from the C-CPG content given in Figure 3, together with their frequency count and UMLS semantic types.

<directive.benefit source="explicit">

Urinary tract infections (UTIs) are important be cause they cause acute morbidity and may result in long term medical problems, including hypertension and reduced renal function. Management of children with UTI involves repeated patient visits, use of antimicrobials, exposure to radiation, and cost. Accurate diagnosis is extremely important for two rea sons: to permit identification, treatment, and evaluation of the children who are at risk for kidney damage and to avoid unnecessary treatment and evaluation of children who are not at risk, for whom interventions are costly and potentially harmful but provide no benefit. Infants and young children with UTI are of particular concern because the risk of renal damage is greatest in this age group and be cause the diagnosis is frequently challenging: the clinical presentation tends to be nonspecific and valid urine specimens cannot be obtained without invasive methods (suprapubic aspiration [SPA], transurethral catheterization).
</directive.benefit>

Figure 3: Exemplar CPG content.

Table 1: The MeSH terms derived from the C-CPG content given in figure 3. Also, shown is the frequency of each term, its UMLS semantic type and its potential as a trigger word for a particular query type.

No	MeSH Term	Frequency	UMLS Semantic Type	Trigger Word For Query Type
1	Urinary Tract Infections	1	Disease or Syndrome	No
2	Urinary Tract	1	Body System, Body Part, Organ, or Organ Component	No
3	Urine	2	Body Substance, Functional Concept	No
4	Infections	1	Disease or Syndrome	No
5	Causes	3	Functional Concept	**Yes (Etiology)**
6	Morbidity	1	Quantitative Concept, Intellectual Product	**Yes (Prognosis)**
7	Hypertension	1	Disease or Syndrome, Sign or Symptom	No
8	Kidney	3	Body Part, Organ, or Organ Component	No
9	Function	1	Biomedical Occupation or Discipline, Physiologic Function	No
10	Children	4	Age Group	No
11	Patient	1	Patient or Disabled Group	No
12	Use	1	Quantitative Concept	No
13	Radiation	1	Natural Phenomenon or Process	No
14	Cost	1	Quantitative Concept	No
15	Extremity	1	Body Location or Region	No
16	Diagnosis	2	Diagnostic Procedure, Functional Concept	**Yes (Diagnosis)**
17	Rat	1	Mammal	No
18	Sons	1	Family Group	No
19	Permit	1	Manufactured Object, Regulation or Law	No
20	Identification	1	Mental Process	No
21	Treatment	2	Therapeutic or Preventive Procedure	**Yes (Therapy)**
22	Evaluation	2	Intellectual Product, Research Activity	No

23	Risk	3	Qualitative Concept, Quantitative Concept	**Yes (Etiology)**
24	DAME	2	Amino Acid, Peptide, or Protein	No
25	Infants	1	Age Group	No
26	Age Group	1	Age Group	No
27	Agar	1	Indicator, Reagent, or Diagnostic Aid, Pharmacologic Substance	No
28	Ageism	1	Individual Behavior, Social Behavior	No
29	Frequency	1	Temporal Concept	No
30	Methods	1	Intellectual Product, Organism Attribute, Intellectual Product	No
31	Aspiration	1	Organ or Tissue Function	No
32	Catheterization	1	Therapeutic or Preventive Procedure	**Yes (Therapy)**

3.3. Search Query Generator

The literature search query builds on the MeSH terms provided by the C-CPG parser, and uses our multi-level literature search methodology to generate the most optimum literature search query.

3.3.1. Query Categorization: Determining Query Type

At the first step, we categorize the query into the following query types: *therapy*, d*iagnosis, etiology and prognosis*. Our query classification strategy is as follows:

1. Check whether the semantic type of the MeSH term is the trigger word for a query type.
2. If yes, increment the count for the query type by the frequency of the MeSH term.
3. Remove the MeSH term that has served as trigger word from the list of candidate MeSH terms.
4. Perform the above for all the candidate MeSH terms.
5. Finally, the query type with the highest count is deemed to be the winner and hence the most appropriate query type for the given set of candidate MeSH terms.

In case, two query types have the same count then the search query is assigned both the winning query types. However, if more than two query types have the same count then the search query is deemed as ‘generic’ and no query type is assigned to it because the query classification strategy has not been able to convincingly determine the query type.

Continuing with the exemplar C-CPG content shown in figure 3 and the MeSH terms derived (see table 1) we can now determine the query type. From table 1, the count for the four different query types is:

[Therapy:3, Diagnosis:2, **Etiology:6**, Prognosis:1]

Given that the query type *etiology* has the highest count (i.e. 6) we determine that the most appropriate query type for the selected C-CPG content is ***etiology***. As per our approach, the trigger words for the designated query type are subsequently removed from the list of candidate query terms. Table 2 shows the remaining candidate query terms, and it may be noted that 6 medical terms—i.e. those medical terms that served as a trigger word for a query type—were removed from the original list of MeSH terms.

Table 2: The remaining MeSH terms after the query type classification step. These terms are subsequently used to determine the age type of the search query.

No	MeSH Term	Frequency	UMLS Semantic Type	Trigger Word For Age Type
1	Urinary Tract Infections	1	Disease or Syndrome	No
2	Urinary Tract	1	Body System, Body Part, Organ, or Organ Component	No
3	Urine	2	Body Substance, Functional Concept	No
4	Infections	1	Disease or Syndrome	No
5	Hypertension	1	Disease or Syndrome, Sign or Symptom	No
6	Kidney	3	Body Part, Organ, or Organ Component	No
7	function	1	Biomedical Occupation or Discipline, Physiologic Function	No
8	Children	4	Age Group	**Yes (Child)**
9	Patient	1	Patient or Disabled Group	No
10	Use	1	Quantitative Concept	No
11	Radiation	1	Natural Phenomenon or Process	No
12	Cost	1	Quantitative Concept	No
13	Extremity	1	Body Location or Region	No
14	Rat	1	Mammal	No
15	Sons	1	Family Group	No
16	Permit	1	Manufactured Object, Regulation or Law	No
17	Identification	1	Mental Process	No
18	Evaluation	2	Intellectual Product, Research Activity	No
19	DAME	2	Amino Acid, Peptide, or Protein	No
20	Infants	1	Age Group	**Yes (Infant)**
21	Age Group	1	Age Group	No
22	Agar	1	Indicator, Reagent, or Diagnostic Aid, Pharmacologic Substance	No
23	Ageism	1	Individual Behavior, Social Behavior	No
24	Frequency	1	Temporal Concept	No
25	Methods	1	Intellectual Product, Organism Attribute, Intellectual Product	No
26	Aspiration	1	Organ or Tissue Function	No

3.3.2. Query Categorization: Determining Age Group

At the second step of the query categorization stage, we categorize the query along the age dimension using the same strategy as used earlier, however here we search for trigger words for the age dimension. Given the MeSH terms in table 2, the count for the four different age-groups is:

[Infant: 1, **Child:4,** Adult:0, Elderly:0]

Hence, we classify the query to target the *child* age-group. The candidate MeSH terms are reduced to 24 and are listed in table 3.

3.3.3. Term Filtering

In the second stage of query generation we build on the results of the previous query cauterization stage to generate a more optimal search query. As mentioned earlier, the term filtering stage attempts to further reduce the dimension of the search query by removing insignificant candidate MeSH terms.

From a practical perspective term filtering is introduced because it was noted that a search query comprising more than 15 search terms is not handled well by MEDLINE. A search query with too many search terms is considered too specific by MEDLINE and as a consequence no medical articles are retrieved. It is hard to decide the optimal length of the search query. For our purposes, based on empirical results, we decided to set the default length of the search query to six (6) terms. If after the term filtering stage the length of the search query is greater than the pre-defined length then we remove terms on a term frequency basis—i.e. terms with a smaller frequency are removed. This process is characterized as *query rollback* and it involves the removal of one term (with the lowest frequency) at a time and then the submission of the relatively shortened search query to PubMed; if meaningful results are obtained then the search process is concluded, else the next lowest frequency term is removed and the process is repeated. We will like to point out that repetitive submissions of ad hoc search queries is not a practical solution considering the search submission constraints imposed by PubMed—i.e. PubMed requires that MEDLINE be queried only once every three seconds. This implies that BiRD needs to wait for at least three seconds before presenting the next search query.

Term filtering has been implemented as three term filters (details are described earlier). Table 3 shows the outcome of the term filtering stage given the list of candidate MeSH terms retained from the query classification stage (shown in table 2).

Table 3: The outcome of the term filtering stage. The last column indicates the filter number responsible for removing the term.

No	MeSH Term	Frequency	UMLS Semantic Type	Removed By Filter #
1	Urinary Tract Infections	1	Disease or Syndrome	Not Filtered
2	Urinary Tract	1	Body System, Body Part, Organ, or Organ Component	**3**
3	Urine	2	Body Substance, Functional Concept	**2**
4	Infections	1	Disease or Syndrome	**3**
5	Hypertension	1	Disease or Syndrome, Sign or Symptom	Not Filtered
6	Kidney	3	Body Part, Organ, or Organ Component	Not Filtered
7	Function	1	Biomedical Occupation or Discipline, Physiologic Function	**3**
8	Patient	1	Patient or Disabled Group	Not Filtered
9	Use	1	Quantitative Concept	**1**
10	Radiation	1	Natural Phenomenon or Process	Not Filtered
11	Cost	1	Quantitative Concept	**1**
12	Extremity	1	Body Location or Region	Not Filtered
13	Rat	1	Mammal	**1**
14	Sons	1	Family Group	**1**
15	Permit	1	Manufactured Object, Regulation or Law	**2**
16	Identification	1	Mental Process	**3**
17	Evaluation	2	Intellectual Product, Research Activity	**2**
18	DAME	2	Amino Acid, Peptide, or Protein	**3**
19	Age Group	1	Age Group	**3**
20	Agar	1	Indicator, Reagent, or Diagnostic Aid, Pharmacologic Substance	**3**
21	Ageism	1	Individual Behavior, Social Behavior	**3**
22	Frequency	1	Temporal Concept	**1**
23	Methods	1	Intellectual Product, Organism Attribute, Intellectual Product	**2**
24	Aspiration	1	Organ or Tissue Function	**3**

At the conclusion of both query generation stages the following output is provided to the next module.

- Query type = Etiology
- Age-Group = Child
- Search Query: {Kidney:3, Hypertension:1, Urinary Tract Infections:1, Patient:1, Radiation:1, Extremity:1} (term:frequency)

3.4. PubMed Query Facilitator

This module is responsible for submitting the search query to PubMed and then collecting the results of the query. PubMed's E-utilities are incorporated within BiRD to facilitate the process. PubMed's E-utilities allow scripts or servlets to easily query the MEDLINE database.

The query rollback component is implemented within this module and is called into action only when no results are obtained from the search query.

3.5. Result Visualizer

This module is responsible for incorporating the search results within the GEM-encoded C-CPG. Recall that BiRD offers the functionality to select multiple C-CPG segments, where each selected C-CPG is treated as a separate query and the search results are therefore the search results need to be associated with the respective C-CPG section. The result visualizer module collects the individual search results—i.e. list of medical articles—for each selected C-CPG and embeds the search results within the C-CPG next to the respective C-CPG section (as shown in figure 4).

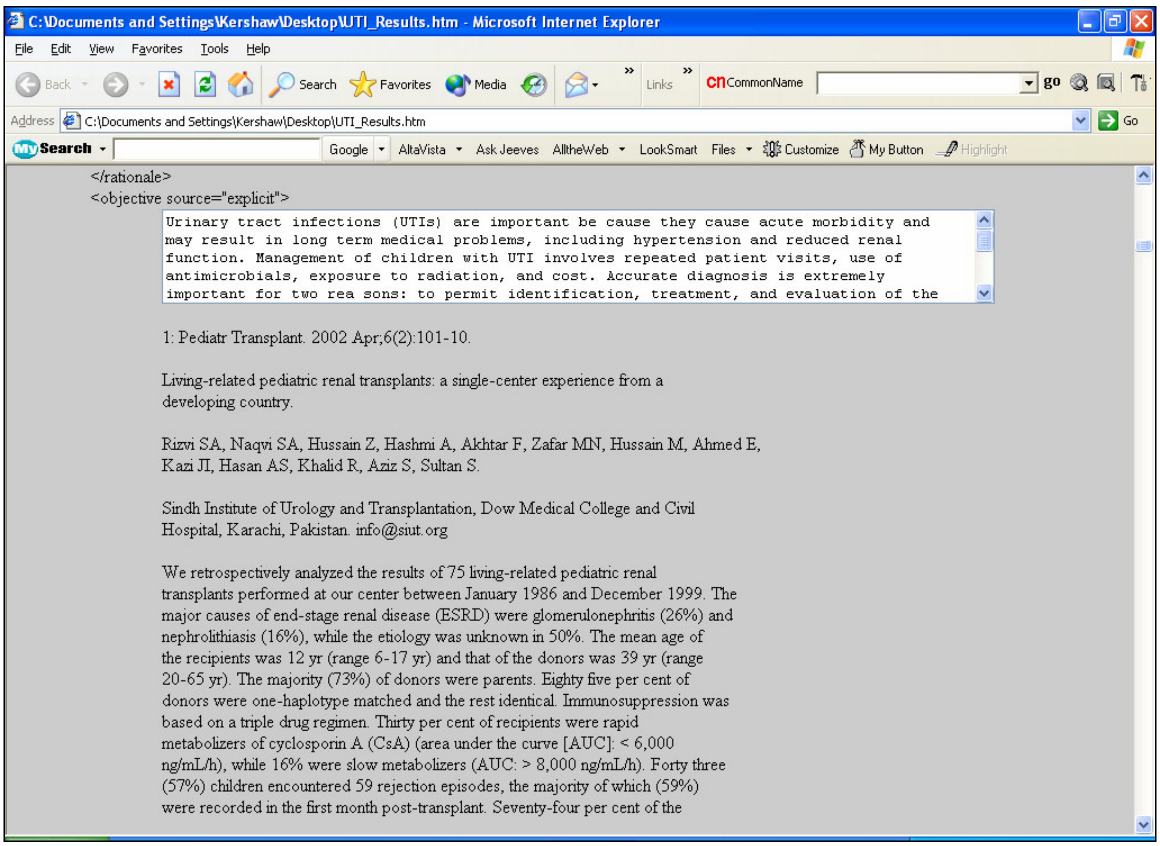

Figure 4: Screen shot of the search results embedded within the GEM C-CPG below the CPG content that was the basis of the search query.

4. Evaluation

For evaluation purposes, the list of medical articles retrieved by BiRD, in response to the above-mentioned query-type, age-group and MeSH terms constituting the search query, is shown in figure 5.

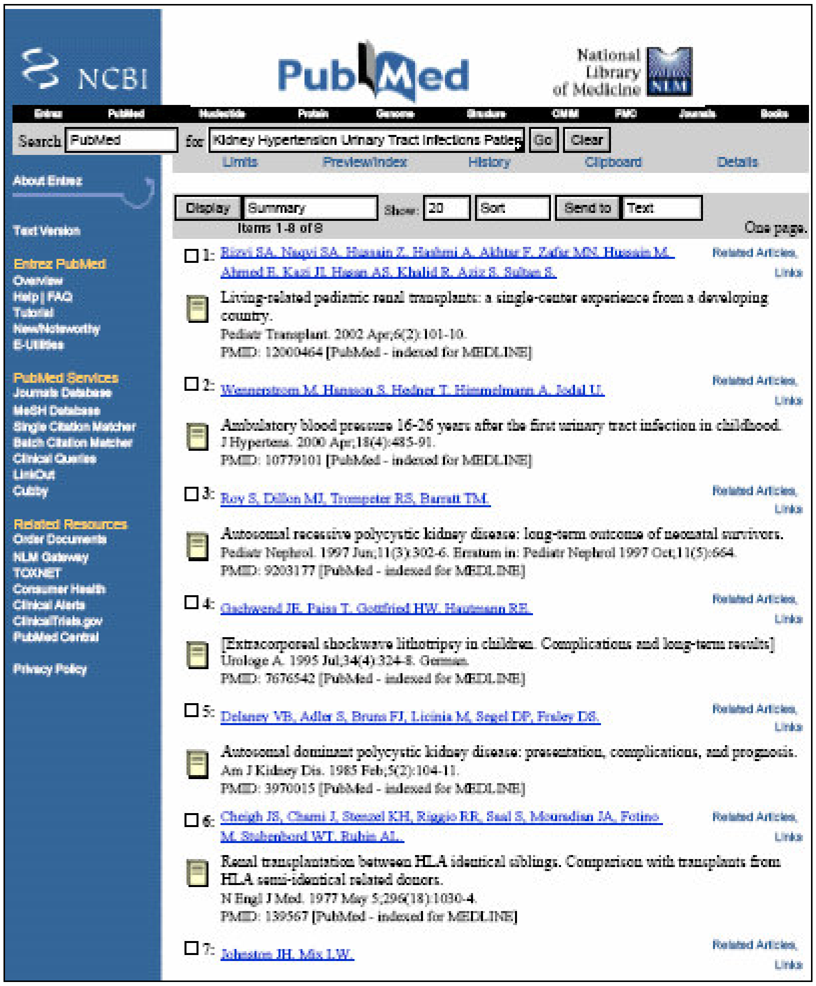

Figure 5: The PuBMed output for the given search query.

The complete list of 8 articles is given below.

1: Rizvi SA, Naqvi SA, Hussain Z, Hashmi A, Akhtar F, Zafar MN, Hussain M, Ahmed E, Kazi JI, Hasan AS, Khalid R, Aziz S, Sultan. Living-related pediatric renal transplants: a single-center experience from a developing country. Pediatr Transplant. 2002 Apr;6(2):101-10.

2: Wennerstrom M, Hansson S, Hedner T, Himmelmann A, Jodal U. Ambulatory blood pressure 16-26 years after the first urinary tract infection in childhood. J Hypertens. 2000 Apr;18(4):485-91.

3: Roy S, Dillon MJ, Trompeter RS, Barratt TM. Autosomal recessive polycystic kidney disease: long-term outcome of neonatal survivors. Pediatr Nephrol. 1997 Jun;11(3):302-6. Erratum in: Pediatr Nephrol 1997 Oct;11(5):664.

4: Gschwend JE, Paiss T, Gottfried HW, Hautmann RE.Extracorporeal shockwave lithotripsy in children. Complications and long-term results. Urologe A. 1995 Jul;34(4):324-8.

5: Delaney VB, Adler S, Bruns FJ, Licinia M, Segel DP, Fraley DS. Autosomal dominant polycystic kidney disease: presentation, complications, and prognosis. Am J Kidney Dis. 1985 Feb;5(2):104-11.

6: Cheigh JS, Chami J, Stenzel KH, Riggio RR, Saal S, Mouradian JA, Fotino M, Stubenbord WT, Rubin AL. Renal transplantation between HLA identical siblings. Comparison with transplants from HLA semi-identical related donors.N Engl J Med. 1977 May 5;296(18):1030-4.

7: Johnston JH, Mix LW. The Ask-Upmark kidney: a form of ascending pyelonephritis? Br J Urol. 1976 Dec;48(6):393-8.

8: Gower PE. A prospective study of patients with radiological pyelonephritis, papillary necrosis and obstructive atrophy. Q J Med. 1976 Apr;45(178):315-49.

We believe that a subjective evaluation of the search results is required to ascertain the efficacy of the search strategy. At this stage, we are currently involved in a number of evaluation studies involving medical practitioners.

We evaluated the impact of the query classification stage on the entire UTI C-CPG mentioned above. In total, the UTI C-CPG (which was encoded in GEM) was divided into 142 sections. Each section was treated as a potential search query content thus realizing 142 search queries. Each of the 142 C-CPG sections was then passed through the MeSH term generator and a potential search query comprising the identified MeSH terms were generated. We divided the length of the search query into four query-size ranges and recorded the number of queries that belonged to each query size range (as shown in table 4). Next, we passed the 142 search queries through the search query generator, recorded the length of the query and tabulated the number of queries in each query-size range (as shown in table xx). The objective of the evaluation experiment was to determine whether there was a reduction in the number of queries in each query-size range after passing through the query categorization algorithm. It may be noted that for the higher query-size ranges the number of queries in each range decreased, most significantly the number of queries in the query-size range 16+ reduced from 10 to 5. These results indicate that the query categorization algorithm is effective, more so for queries comprising a larger number of query terms.

Table 4: A comparative analysis of the query classification algorithm

Query-size Range	**Before Query Classification**	**After Query Classification**
0 – 4	99	111
5 -9	19	15
10 – 15	14	11
16+	10	5
Total	**142**	**142**

5. Concluding Remarks

In this paper we have presented a medical literature search strategy that aims to generate an optimal search query for retrieving medical literature from MEDLINE. The featured search strategy builds a search query in an objective manner, as opposed to the traditional subjective manner of query specification by users, by using the original medical content as the basis for deriving the search query. We believe that such an approach would have implications in (a) easing the query specification process for health practitioners; and (b) improving the quality of best evidence retrieved.

Our work is a first step in realizing the need for supplementing C-CPG with best evidence—evidence that is autonomously sourced from medical literature archives. We have presented a working *proof-of-concept* that clearly demonstrates how to link best evidence vis-à-vis medical articles, to existing C-CPG. We believe that this concept will be appealing to health practitioners and will improve the acceptance and efficacy of C-CPG in clinical settings. Our novel BiRD system, incorporating our multi-level search strategy—provides the basis for a testbed to extend the functionality of linking medical literature with other C-CPG formalisms. The work is in progress and within a month evaluation results will be available. However, we have identified a number of issues that need to be addressed in the realm of future work, for instance investigating the possibility to use the semantically-explicit tags in the GEM document to assist in query-type categorization. Furthermore, in the light of empirical evaluations we intend to fine tune some of the initial assumptions regarding the filterable semantic types, optimal query length and query rollback scheme.

We will like to point out that our query classification results are positive and vindicate our search query generation model. Indeed, the retrieval results are directly related to the query categorization accuracy. Having said that, query types are not necessarily formal, rather abstractions of medical queries typically posed by health practitioners. To formalize an objective set of query types is a major research endeavor and a task for future research.

References

[1] Field MJ, Lohr KN (Eds). Guidelines for clinical practice: from development to use. Institute of Medicine, Washington, D.C. National Academy Press, 1992.

[2] Institute of Medicine. Guidelines for clinical practice: from development to use. Washington, DC, National Academy Press, 1992.

[3] Shiffman RN, Liaw Y, Brandt CA, Corb GJ. Computer-based guideline implementation systems: a systematic review of functionality and effectiveness. J Am Med Inform Assoc 1999;6:104–114.

[4] Davis DA, Taylor-Vaisey A. Translating guidelines into practice: A systematic review of theoretic concepts, practical experience and research evidence in the adoption of clinical practice guidelines. Can Med Assoc J 1997; 157:408–416.

[5] Tierney WM, Overhage JM, McDonald CJ, et al. Computerizing guidelines to improve care and patient outcomes: The example of heart failure. J Am Med Inform Assoc 1995; 2:316–322.

[6] Lobach DF, Hammond WE. Computerized decision support based on a clinical practice guideline improves compliance with care standards. Am J Med 1997; 102:89–98.

[7] Maviglia SM, Zielstroff RD, Paterno M, Teich JM, Bates DW, Kuperman GJ. Automating complex guidelines for chronic disease: lessons learned. J Am Med Inform Assoc. 2003;10:154–165.

[8] Chueh H, Barnett GO. "Just-in-time" clinical information. Acad Med 1997; 72:512–517.

[9] Shiffman RN, Karras BT, Nath S, et al. GEM: A proposal for a more comprehensive guideline document model using XML. J Am Med Inform Assoc 2000; 7:488–498.

[10] Haynes B, Wilczynski, McKibbon A, Walker C, Sinclair J. Developing optimal search strategies for detecting clinically sound studies in MEDLINE. J Am Med Infrm Assoc. 1994;1:447-458.

[11] Cimino J, Mendonca E. Automated knowledge extraction from MEDLINE citations. Proc AMIA Symp. 2000;:575-9.