Towards Patient-Related Information Needs

Loes Braun[a], Floris Wiesman[b], Jaap van den Herik[a], Arie Hasman[b], Erik Korsten[c]

[a] *Institute for Knowledge and Agent Technology, University Maastricht, The Netherlands*
[b] *AMC, University of Amsterdam, The Netherlands*
[c] *Catharina-ziekenhuis, Eindhoven, The Netherlands*

Abstract

The quality of health care depends, among others, on the quality of a physician's domain knowledge. Since it is impossible to keep up with all new findings and developments, physicians usually have gaps in their domain knowledge. To handle exceptional cases, access to the full range of medical literature is required. The specific literature needed for appropriate treatment of the patient is described by a physician's information need. Physicians are often unaware of their information needs. To support them, this paper[1] aims at presenting a first step towards automatically formulating patient-related information needs. We start investigating how we can model a physician's information needs in general. Then we propose an approach to instantiate the model into a representation of a physician's information needs using the patient data as stored in a medical record. Our experiments show that this approach is feasible. Since the number of formulated patient-related information needs is rather high, we propose the use of filters. Future research will focus on the combination of personalization and filtering. It is expected that the resultant set of information needs will have a manageable size and contributes to the quality of health care.

Keywords:
Medical Records Systems, Computerized; Information Storage and Retrieval; Quality of Health Care

1. Introduction

We start with an example that precisely illustrates the need for knowledge of patient-related literature.

An 84-year-old woman was brought into the emergency department of a hospital, suffering from dyspnea and loss of consciousness. Five days earlier she had visited her general practitioner who diagnosed her with suspected respiratory tract infection and prescribed a drug called Clarithromycin. However, instead of improving, her condition worsened. In the hospital the diagnosis pneumonia was considered and she was treated accordingly, but without any effect. Upon her family's request, the patient was not admitted to the intensive care unit and she died one day after she was admitted to the hospital. Surprisingly, an autopsy revealed that the cause of death was not pneumonia, but a case of severe acute pancreatitis. The autopsy also revealed that the most plausible cause for the pancreatitis was the use of Clarithromycin, since pancreatitis is a (rare) side effect of the use of Clarithromycin [1].

Since the incidence of Clarithromycin-induced pancreatitis is quite low, it is

[1] This research is part of the MIA project (Medical Information Agent), which is funded by NWO (grant number 634.000.021).

understandable (but still undesirable) that the physician in the example above was not aware of this possible side effect. If the physician had performed a literature search in Medline on the side effects of Clarithromycin, he[2] probably would have found an article by Leibovitch, Levy, and Shoenfeld [2], in which another case of Clarithromycin-induced pancreatitis is discussed. If he had read this article, he probably would have ordered additional diagnostic tests to exclude pancreatitis (e.g., blood amylase) and he could have started the appropriate treatment immediately.

We define an *information need* as a formulation of missing information needed to perform a particular task. In our example the physician's information need was *What are the side effects of Clarithomycin?* However, the physician was not aware of his information need. Therefore, we call the information need *implicit*, as opposed to an *explicit* information need of which one *is* aware. Since the physician's information need was implicit, he had no incentive to search for information on the topic. Hence, our conclusion from the example is that information needs should be made explicit automatically in order to perform an appropriate literature search.

The example above clearly illustrates that (automatic) retrieval of relevant, patient-related literature is vital to the quality of care (cf. [3]). Various articles discuss information-retrieval (IR) systems that provide such literature (e.g., [4, 5, 6, 7]); our research roughly follows the contents of these articles. However, in our opinion the overall shortcoming of the systems mentioned in the articles is that the degree of necessary interaction with the systems is too high. This is especially true in the area of making information needs explicit. Therefore, our main research objectives are (1) to investigate to what extent a physician's implicit information needs can be made explicit automatically, and (2) to implement our approach together with some filters into a computer system supporting physicians in their daily work.

Section 2 describes how we determine a physician's information needs and how we model these needs. Section 3 presents our approach to formulate patient-related information needs (i.e., based on the patient and the physician's current activities with respect to the patient). In section 4 experiments and results are shown and briefly discussed. Section 5 provides our conclusions and directions for future research.

2. Modelling a Physician's Information Needs

Our approach to make a physician's information needs explicit is to anticipate them. As a starting point for this process, we need a set of a physician's potential information needs. However, such a set can never be complete, since it is impossible to capture all of a physician's information needs. Moreover, a physician generates new information needs over time, which should be added to the set. This is hard to facilitate.

One solution is to build a *model* of a physician's information needs. As long as the model represents information needs on a more abstract level it can be considered complete, meanwhile anticipating future information needs. Modelling a physician's information needs involves two steps described below: (1) identifying a physician's information needs (subsection 2.1) and (2) abstracting the identified information needs (subsection 2.2).

2.1 Identifying a Physician's Information Needs

To identify a physician's information needs, we used two methods, viz. (1) a literature survey and (2) interviews. Both identification methods are briefly described below. Table 1 summarizes the sources, the identification domains, and the number of information needs identified.

[2] For brevity we will use the pronoun 'he' ('his') where 'he or she' ('his or her') is meant.

Table 1 - Number of information needs identified by a literature survey and interviews.

Identification method	**Source**	**Identification domain**	**# INs identified**
Literature survey	[8]	Outpatient care, inpatient care, internal medicine	16
	[9]	General practice, cardiology, pulmonology, allergology	77
	[10]	Family care	10
	[11]	Primary care	16
	[12]	Various	32
	[13]	Various	10
	[14]	Surgical care	2
	[15]	Primary care	8
Interviews		Anaesthesiology	2
		Cardiology	1
		Neurology	0
		Pulmonology	3
		Surgery	3

In our literature survey, we searched for articles presenting information needs that are general, i.e., not specific for a particular group of physicians or for a particular geographical area. We found only eight such articles [8-15]. This set of articles covered a large number of medical domains from which the information needs were identified. In total we arrived at 171 information needs.

To obtain a set of information needs that is as diverse as possible, we succeeded in interviewing five physicians in five different medical specialisms: (1) anaesthesiology, (2) cardiology, (3) neurology, (4) pulmonology, and (5) surgery. The physicians were interrogated by means of an interview scheme composed in advance. This led to 9 additional information needs.[3]

2.2 Abstracting the Identified Information Needs

The identified information needs are highly context-dependent, which may render them useless in another (different) context. To reduce context-dependency, we abstracted the information needs, so as to make them context-independent. For the abstraction we used an approach similar to the one used by Ely, Osheroff, and Ebell [10]. We replaced each medical concept in the information needs by its semantic type, which is a high-level description of the medical concept (e.g., the concept 'Pneumonia' has the semantic type DISEASE OR SYNDROME). We obtained the semantic types of the concepts from the Semantic Network of the Unified Medical Language System (UMLS) that comprises 135 types [16].

The abstraction resulted in a general class of information needs, called *information-need templates*. Some information needs resulted in the same information-need template. For example, *Does Morphine cause rash?* and *Does Clarithromycin cause high blood pressure?* both resulted in the information-need template *Does [CHEMICAL] cause [SIGN OR SYMPTOM]?* To obtain a proper *set* of information-need templates, we removed all doubles. Currently, the set comprises 167 information-need templates.

3. Instantiating Templates into Patient-Related Information Needs

To represent a patient-related information need, an information-need template has to be instantiated with a patient's medical data. The data are acquired from the EPR (Electronic

3 Since we have to search English literature and several information needs were in Dutch, we translated the Dutch information needs into English.

Patient Record) of the specific patient. In our research we employed the *Intensive Care Information System*,[4] used at the Intensive Care Unit of the Catharina-ziekenhuis in Eindhoven.

Our approach of converting an information-need template into the representation of a patient-related information need comprises three steps, viz. (1) select EPR-queries that indicate the appropriate patient data[5] in the EPR, (2) execute the selected EPR-queries: the desired data are extracted from the EPR, and (3) instantiate the information-need template with the results of the executed queries.

In the first step, we start determining which semantic types occur in the information-need template. To convert the information-need template into an information need, each of these semantic types has to be instantiated with patient data. Consequently, an EPR-query has to be selected for each semantic type in the information-need template. The EPR-queries are selected from a list of EPR-queries, formulated in advance. Each EPR-query in this list specifies how to find the patient data associated with the corresponding semantic type. Assume we have the template *What are the side effects of [CHEMICAL]?* Based on (1) the semantic type CHEMICAL, (2) the information structure of our EPR, and (3) the patient number of the specific patient, the following EPR-query is selected (the names of the database tables are in Dutch) *SELECT Medicijn FROM Medicatie WHERE PatientNummer=1234567890*. To facilitate easy adaptation to other EPR-systems, all potential EPR-queries for a specific EPR-system are specified in a model, which is runtime consulted by the system and can be easily reformulated.

The second step is to execute the selected queries to extract the desired patient data from the EPR. The actual query-execution process is handled by the database itself. Each result that an EPR-query returns for a semantic type is called an *active concept* of that specific semantic type. Assume that our patient is taking three different medications. Then, our EPR-query has three results and consequently, the semantic type CHEMICAL has three active concepts, e.g., (1) *Clarithromycine*, (2) *Amoxi/Clavulaan*, and (3) *Furosemide-iv*. Since all terms from the EPR (and consequently also all active concepts) are in Dutch, we mapped them manually to UMLS concepts, which are then translated into English by means of the UMLS Metathesaurus.

The third step is to instantiate the information-need template with the data obtained from the EPR (active concepts of the semantic types). We call an information-need template *applicable* (i.e., it can be instantiated with patient data) if each semantic type within the information-need template has one or more active concepts. If the template is applicable, it is instantiated by systematically replacing each semantic type by one of its active concepts, until all possible combinations are used. The total number of resulting information needs is the product of the numbers of active concepts of all semantic types in the template. For the three active concepts of the semantic type CHEMICAL, our information-need template is instantiated three times, viz. (1) *What are the side effects of Clarithromycin?*, (2) *What are the side effects of Amoxicillin-Clavulanic Acid?*, and (3) *What are the side effects of Furosemide*? If a literature search were conducted, based on the above information needs, patient-related literature would be found. The approach described above was implemented in a computer system.

---

[4] Intensive Care Informatie Systeem, Version 2.8. INAD Computers & Software B.V. Eindhoven, Werkgroep ICIS Afd. Intensive Care, Dienst Informatie Voorziening, Catharina-ziekenhuis Eindhoven.

[5] All patient data used in the examples of section 3 are fictitious.

4. Experiments and Results

To establish the feasibility of our approach for instantiating information-need templates, we let our system formulate patient-related information needs based on the EPRs of 82 patients. Each EPR contained information about *all* hospital adm sions (in the hospital under consideration) of a patient. After each separate data entry, new information needs were formulated based on the added data (possibly in combination with already available data). We used our complete set of 167 information-need templates. For each patient, we calculated the average number of information needs formulated per data entry. Each patient was placed into one of four categories, based on the average number of information needs formulated, viz. (i) no information needs, (ii) a manageable number of information needs (1-100), (iii) a hardly manageable number of information needs (101-1000), and (iv) and

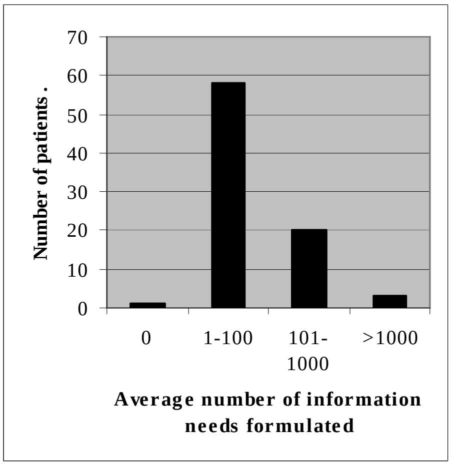

Figure 1 - Number of patients for which a specific number of information needs is formulated.

unmanageable number of information needs (>1000).

Figure 1 shows how many patients were placed in each category. As can be seen in figure 1, in 71% of the cases a manageable number of information needs is formulated. However, for 28% of the cases the number of information needs formulated is hardly manageable or even unmanageable. If a literature search were conducted for all these information needs, the set of retrieved literature would be unmanageably high for these cases. Since we do not want to overload physicians with literature, the number of information needs should be reduced by filtering. In the ideal situation, all patients would be in category 1-100 information needs.

To reduce the number of formulated information needs two filters can be used. The first one is the *stage of the medical process*. Not all information needs might apply to the current stage of the medical process. For example, when a physician has already selected chemotherapy as the appropriate treatment for a lung-cancer patient, he is assumed not to have information needs concerning the selection of a treatment, such as *What is the treatment for lung cancer for this patient?* Yet, he might have information needs concerning the execution of the selected treatment, such as *How high is the dose of chemotherapy for lung cancer for this patient?* The second filter is the *specialism of the physician*. Since a physician's information needs are probably connected to his specialism, we might ignore several information needs, because they are not linked to the physician's specialism. Ignoring the information needs is solely based on the patient data with which the corresponding information-need templates were instantiated. As information-need templates contain no patient data, templates cannot be ignored in advance. A series of future experiments may clarify to what extent the two filters are appropriate for a reduction of the number of information needs.

5. Conclusions

The cooperation with physicians, described in section 2, showed that we succeeded in identifying a physician's potential information needs and in modelling them into 167 information-need templates by using 135 semantic types.

In section 3 we designed an approach to convert information-need templates into patient-related information needs by taking patient data into account. From the experiments, we may conclude that our approach is adequate and can be generalized to other EPR-systems, as long as they use a suitable information structure. The number of automatically formulated, patient-related information needs per patient is still high (section 4), but may be reduced by taking the *stage of the medical process* and the *specialism of the physician* into account.

When using the automatically formulated information needs as a starting point for literature retrieval, patient-related literature can be provided to the physician automatically, thereby potentially contributing to an improvement in the quality of health care.

6. References

[1] Schouwenberg BJJW and Deinum J. Acute pancreatitis after a course of Clarithromycin. *The Netherlands Journal of Medicine* 2003: 61(7) pp. 266–267.

[2] Leibovitch L, Levy Y, and Shoenfeld, Y. Pancreatitis induced by Clarithromycin. *Ann Intern Med* 1996: 125(8) pp. 701.

[3] Gamble S. Hospital libraries enhance patient care and save money. *Journal of the Alberta Association of Library Technicians* 1996: 23(2) pp. 10–12.

[4] Miller RA, Jamnback L, Giuse NB, and Masarie FE. Extending the capabilities of diagnostic decision support programs through links to bibliographic searching: Addition of 'canned MeSH logic' to the Quick Medical Reference (QMR) program for use with Grateful Med. In Clayton P, ed. *Proc Annu Symp Comp App Med Care*. McGraw-Hill Inc., 1991; pp. 150–155.

[5] Rada R, Barlow J, Bijstra D, Potharst J, de Vries Robbé P, and Zanstra P. OAR: Open architecture for reasoning applied to connection patient records to medical literature. In Noothoven van Goor J and Christensen JP, eds. *Advances in Medical Informatics*. Amsterdam: IOS Press, 1992; pp. 287–294.

[6] Van Mulligen EM. UMLS-based access to CPR data. *Int J Med Inf* 1999: 53(2–3) pp. 125–131.

[7] Cimino JJ, Johnson SB, Aguirre A, Roderer N, and Clayton PD. The Medline Button. *Proc Annu Symp Comput Appl Med Care* 1992; pp. 81-85.

[8] Cucina RJ, Shah MK, Berrios DC, and Fagan LM. Empirical formulation of a generic query set for clinical information retrieval systems. In Patel V, Rogers R, and Haux R, eds. *Proc MedInfo2001*. Amsterdam: IOS Press, 2001; pp. 181–185.

[9] De Vries Robbé PF, Beckers WPA, and Zanstra PE. *MEDES. Het prototype*. Groningen: Onderzoeksgroep Medische Informatie- en Besliskunde, Academisch Ziekenhuis Groningen, 1988. (In Dutch).

[10] Ely JW, Osheroff JA, and Ebell MH. Analysis of questions asked by family doctors regarding patient care. *BMJ* 1999: 319(7206) pp. 358–361.

[11] Gorman PN. Information needs of physicians. *J Am Soc Inf Sci* 1995: 46(10) pp. 729–736.

[12] Grundmeijer HGLM, Reenders K, and Rutten GEHM, eds. *Het Geneeskundig Proces. Van Klacht naar Therapie*. Maarssen: Elsevier Gezondheidszorg, 1999. (In Dutch).

[13] Jerome RN, Giuse NB, Gish KW, Sathe NA, and Dietrich MS. Information needs of clinical teams: analysis of questions received by the clinical informatics consult service. *Bull Med Lib Assoc* 2001: 89(2) pp. 177–185.

[14] Reddy MC, Pratt W, Dourish P, and Shabot MM. Asking questions: Information needs in a surgical intensive care unit. In Kohane IS, ed. *Proc AMIA'02.* Philadelphia: Hanley and Belfus Inc., 2002: pp. 647–651, 2002.

[15] Smith R. What clinical information do doctors need? *BMJ* 1996: 313(7064) pp. 1062–1068.

[16] U.S. National Library of Medicine. Unified Medical Language System, 2003. http://www.nlm.nih.gov/research/umls/.

Address for correspondence

Drs. L.M.M. Braun, Institute for Knowledge and Agent Technology, University Maastricht, P.O. Box 616, 6200 MD Maastricht, The Netherlands, L.Braun@cs.unimaas.nl.