Eliciting user decision requirements for designing computerized diagnostic support for family physicians

(T Journal accepted manuscript version)

Eliciting user decision requirements for designing computerized diagnostic support for family physicians

Talya Porat¹, PhD, Research Fellow in Human Factors

Olga Kostopoulou*², PhD, Reader in Medical Decision Making

Amanda Woolley¹, PhD

Brendan C Delaney¹,², MD, Professor of Medical Informatics and Decision Making

¹ Department of Primary Care and Public Health Sciences, Faculty of Life Sciences & Medicine, King’s College London, UK

² Department of Surgery and Cancer, Faculty of Medicine, Imperial College London

*Correspondence to:

Dr Olga Kostopoulou
Imperial College London
Division of Surgery, Department of Surgery and Cancer
5th floor Medical School Building
St Mary’s Campus
Norfolk Place
London W2 1PG
UK

E-mail: o.kostopoulou@imperial.ac.uk
T: +44 (0)20 7594 9120
Authors biographies

Talya Porat is a researcher in human-computer interaction, human factors and cognitive engineering, with a special interest in the medical field. She led the work on interface design for the diagnostic support tool.

Olga Kostopoulou is a psychologist studying judgment and decision making, with a focus on diagnostic reasoning, error and cognitive biases. She led the overall project for the development and evaluation of the computerized diagnostic support tool.

Amanda Woolley obtained her PhD in Psychology under the supervision of Olga Kostopoulou. She studied clinical intuition in family medicine from a cognitive psychology perspective.

Brendan C. Delaney is the Scientific Director of TRANSFoRm (transformproject.eu). He is an academic primary care physician with an interest in diagnosis and medical informatics.

ABSTRACT

Despite its 40-year history, computerized diagnostic support is not used in routine clinical practice. As part of a European project to develop computerized diagnostic support for family physicians, we identified user decision requirements and made design recommendations. To this end, we employed multiple data types and sources. All data were elicited from UK family physicians and pertained to consultations with patients, either real or simulated. To elicit user requirements, we conducted in situ observations and interviews with 8 physicians, and performed a hierarchical task analysis of the diagnostic task. We also analysed 34 think-aloud transcripts of 17 family physicians diagnosing detailed patient scenarios on a computer, and 24 interview transcripts of 18 family physicians describing past cases of intuitive diagnoses from their experience. All transcripts were coded using the Situation Assessment Record (SAR) method. We report our methods and results using the Decision-Centered Design framework. Studies employing multiple human factors techniques and data types in order to elicit user requirements are rare. Our approach enabled us to propose
interface design recommendations that go beyond existing “differential diagnosis generators”, aiming to improve physicians’ performance and acceptance of the resulting tool.

**Keywords:** Family physicians; Clinical Decision Support Systems; decision making; diagnostic error; diagnosis; diagnostic support
INTRODUCTION

Family physicians in the UK have a gatekeeping role, controlling access to specialist services. Thus, one of their key challenges is trying to balance the risk of missing a serious disease against unnecessary investigation or referral. Although the prevalence of serious disease is relatively low in primary care, the sheer volume of patient contacts (90% of contacts in the UK healthcare system) means that only very low risks can be tolerated. Data from both major UK medical defense organizations show that diagnostic error is the reason for most patient claims against family physicians (63%-66%) (Silk 2000). It is also the most common reason for malpractice claims in the ambulatory care setting in the USA (59%) (Gandhi, Kachalia et al. 2006).

Family physicians see a relatively large number of patients in short 10-minute consultations, i.e., around 20 patients per clinical session (3.5 hours). They deal with a wide range of disease areas and patients suffering from multiple conditions (Salisbury, Procter et al. 2013). They also deal with non-specific symptoms that could be attributed to a number of causes. These are some of the factors that explain why diagnosis in family medicine can be challenging. In addition, the multiplicity and heterogeneity of tasks that UK family physicians are expected to perform during a 10-minute consultation, such as screening for certain diseases, and health promotion, and the cost considerations in relation to diagnostic tests and specialist referrals, can exert more pressure on family physicians and reduce even further the time that they can dedicate to the important task of diagnosis. In the US, family physicians are under similar pressures, with the average time for consultations declining and the added problem of increasing amounts of paperwork required for administration and billing purposes (Musen, Middleton et al. 2014).

Diagnostic error results from both factors in the healthcare system and clinical judgment (cognitive factors). Cognitive factors are thought to be the most prevalent cause of
diagnostic error (Graber, Franklin et al. 2005). In a large retrospective study of diagnostic adverse events in Dutch hospitals, cognitive factors were found to have played a significant part in 96% of the events and system failures in only 25% (Zwaan, de Bruijne et al. 2010). A US study of closed malpractice claims (patients alleging missed or delayed diagnosis) in the ambulatory setting estimated that cognitive factors (e.g. judgment errors, vigilance and memory lapses, lack of knowledge) were implicated in virtually all diagnostic errors, either alone (in 55% of errors) or in association with patient- and/or system-related factors (Gandhi, Kachalia et al. 2006). The most frequent breakdowns in the diagnostic process were failure to order appropriate diagnostic tests (55%), failure to follow up appropriately (45%), inadequate history taking and physical examination (42%), and incorrect interpretation of diagnostic tests (37%), mostly imaging. It is apparent that failure to gather sufficient and appropriate information was responsible for most errors. In their seminal programme of work, Elstein and colleagues ascertained the importance of diagnostic hypotheses driving both the search for and interpretation of clinical information (Elstein, Shulman et al. 1978) (Elstein, Shulman et al. 1990).

Family physicians report that they diagnose using mostly automatic or semi-automatic strategies, based on rules or the recognition of patterns, rather than any analytical reasoning (Heneghan, Glasziou et al. 2009). These heuristic strategies serve them well in familiar and routine situations because they can help them arrive at the right diagnosis or decision quickly, confidently and efficiently. Nevertheless, heuristics that are adapted to one environment (Gigerenzer, Todd et al. 1999) can easily turn to biases (Kahneman, Slovic et al. 1982), when changes in the environment go unnoticed by the decision maker. For example, in the absence of a back injury, back pain that feels worse when the patient wakes up in the morning is most likely mechanical. However, back pain that wakes up the patient during the night may suggest that it is possibly due to a more serious cause. If physicians have not generated the right hypotheses to account for the presenting symptoms, they may ignore them, explain them away (Kostopoulou, Devereaux-Walsh et al. 2009) or change
Eliciting user decision requirements for designing computerized diagnostic support for family physicians


Computerized diagnostic support systems (referred hitherto as “diagnostic systems” or simply “systems”) have been developed since the early ’70s (de Dombal, Leaper et al. 1972) (Leaper, Horrocks et al. 1972). Studies assessing their effectiveness in improving physician performance have produced mixed results (Garg, Adhikari et al. 2005). Although they may remind physicians of diagnoses that they would otherwise have not considered, this has only been demonstrated in experimental settings (Berner 2009) and their use in clinical practice is very limited.

There are two major technical problems with existing diagnostic systems that constitute significant barriers to their adoption and effective use: lack of integration with the electronic health record (EHR) and lack of consideration of the physician’s diagnostic workflow (El-Kareh, Hasan et al. 2013) (Shibl, Lawley et al. 2013) (Kawamoto, Houlihan et al. 2005). Proper elicitation of user requirements is an essential pre-requisite for the design of any application that seeks to support diagnosis within the clinical encounter, yet this has not been carried out previously. Commercial systems (“differential diagnosis generators”), such as DXplain™, Isabel, and SimulConsult, are stand-alone applications, requiring physicians to switch from their EHR to the system and enter information twice. This presumes that physicians recognise the need for advice and are sufficiently motivated to spend the time entering information and examining system advice. The evidence for either of these assumptions is discouraging (Friedman, Gatti et al. 2005) (Ramnarayan, Winrow et al. 2006). Furthermore, even when physicians decide to consult the system, they will do so after they have collected substantial information from the patient. It follows that system advice based on that information may well be biased by the hypotheses that the physician has already considered: “The system’s advice, and thus its potential value, depends on how
users can convey to the DSS their personal understanding of a case by selectively entering clinical findings…” (Friedman, Elstein et al. 1999) p. 1852.

In a recent experimental study (Kostopoulou, Rosen et al. 2015), family physicians diagnosed a number of patient scenarios on a computer. The study examined two types of automated diagnostic support: one where a list of diagnostic suggestions is provided early on in the clinical encounter, triggered by the reason for encounter and the patient’s risk factors, before physicians start asking questions to test their hypothesis/es; and one where an individualized, shorter list of diagnostic suggestions is provided late in the encounter, based on information that the physician has collected and triggered by the physician entering his/her diagnosis. Both types of support were tested against an unaided control group. The study found that early support significantly improved diagnostic accuracy over control, without lengthening information search and time taken, while late support was no more accurate than control. The effect was replicated in Greece, a European country with an entirely different healthcare system, demonstrating the generalizability of this generic, early intervention to improve diagnostic accuracy (Kostopoulou, Lionis et al.). We adopted the principle of early support in the design of a computerized diagnostic support system for family medicine and carried out a user requirements elicitation process described in this paper.

Elicitation of user requirements is a critical and complex phase in the design and development of information systems. Inappropriate or insufficient elicitation, e.g., based on a single method or limited numbers of participants, can lead to failed system functionality and user adoption (Zowghi and Coulin 2005, Davey and Cope 2008). We aimed to elicit user requirements for the design of a prototype for computerized diagnostic support for family physicians to be developed as part of the EU FP7 TRANSFoRm project (transformproject.eu).
METHODS

We employed a Decision-Centered Design framework (DCD) (Crandall, Klein et al. 2006) (Militello and Klein 2013) one of several frameworks for cognitive systems engineering. DCD advocates focusing on difficult key decisions and non-routine situations, where errors may lead to injury and/or death. It is therefore suitable for designing support for medical diagnosis. DCD uses cognitive task analysis (CTA) methods to identify the key decisions. It then translates these into cognitive requirements. The system design process focuses on these requirements to support decision making in challenging situations, assuming that the routine requirements will be incorporated along the way. DCD aims to ensure that the design addresses cognitive challenges, so that cognitive performance is improved, and the human-computer interface reflects the users’ needs.

The DCD framework includes five phases. Here, we describe the first three phases for the design of the prototype: preparation, knowledge elicitation, and analysis and representation. We are currently engaged in the last two DCD phases, application design and prototype evaluation, hence these are not reported in this paper. Figure 1 shows the first three phases of the DCD framework, with their respective data sources, analyses and outputs.

Preparation: The preparation phase seeks to gather background material about the domain, the nature and range of the tasks involved, and to identify cognitively complex task elements. Preparation started with reviewing an existing hierarchical task analysis (HTA) of the family medicine consultation (Kostopoulou 2006). HTA models tasks as hierarchies of goals and sub-goals, with plans that show how sub-goals should be carried out (Annett and Duncan 1967) (Shepherd 2001). We aimed to refine the parts of the HTA that related to diagnosis. For this purpose, we observed 8 family physicians (5 male, mean 8.6 years in family medicine, SD 6) consulting with their patients, which resulted in the observation of 104 clinical encounters (23.5 hours) in total. A researcher (TP) sat in the consulting room
and unobtrusively observed the clinical interaction, taking notes of the tasks the physicians performed, their workflow and how they used their EHR, i.e., for which tasks and at which stage in the workflow. Notes from all observations were compared, which helped us to focus on the observable behaviours and interactions with the EHR, and refine the existing HTA (Figure 2).

The notes also guided the post-observation interviews of the 8 physicians, which focused on the clinical encounters observed earlier. The aim of these interviews was to confirm the flow and tasks involved in diagnosing a patient, as were observed, and to identify cognitively complex task elements in the diagnostic process. We employed “intensive interviewing” (Legard, Keegan et al. 2003) rather than a structured interview. Intensive interviews are adaptive to the situation of interest and allow the content and order of the questions to vary from one interviewee to another. In addition to specific questions about each clinical encounter observed, we asked physicians to think back to past diagnostic errors and suggest how computerized diagnostic support might have helped to avoid these errors.

Physicians’ answers were documented and examined for important concepts and their relationships. We identified the cognitive aspects and elements (e.g., potential for errors, difficulties, and strategies) of the diagnostic task. For example, in relation to the HTA sub-task “Get familiar with patient’s clinical history”, the physicians described the difficulty in identifying and retrieving from the patient record information that could potentially be critical for diagnosis, e.g., similar previous episodes, co-morbidities, and risk factors, and their strategies of doing so, e.g., filtering according to high priority problems. This information was then used in the Decision Requirements Table.

*Knowledge Elicitation:* The knowledge elicitation phase uses cognitive task analysis (CTA) methods to elicit critical incidents and key components of expert decision making. For this purpose, we used CTA methods to analyse two types of existing verbal data: a) think-aloud
protocols of family physicians diagnosing patient cases presented on a computer and b) interview protocols of family physicians describing past cases of intuitive diagnoses.

The think aloud protocols were collected during an ongoing study by the second author that investigates how family physicians deal with early presentations of cancer. Participating physicians viewed a series of patient cases on computer. After some initial information about the patient and his/her main health complaint, physicians requested further information in order to diagnose. They could take a history and request results of physical examinations and laboratory tests. A researcher provided responses from a predetermined list. Furthermore, participants were instructed to think out loud (Ericsson and Moxley 2011). We used the first 34 think-aloud protocols from this study: 11 pertained to a lung cancer scenario, 11 to a myeloma scenario and 12 to a colorectal cancer scenario. Lung and colorectal cancers are common cancers, while myeloma is rare.

Using data thus obtained provided considerable control, since it enabled us to study challenging diagnostic situations with predetermined difficulty, presented in a standard way to multiple participants; this enabled comparisons between transcripts and identification of diagnostic strategies. In addition, each scenario contained critical information that could be obtained upon the physician’s request, so we could identify omissions in information search, and interpretation errors. The computer program in the original study automatically recorded each physician’s sequential information acquisition. This gave a structure to the verbal reports and provided a validity check. Finally, each scenario had an optimal solution, i.e. depicted a specific diagnosis, against which participants’ accuracy was measured.

The think aloud protocols were elicited in a study where all information was provided in written form, the physicians did not have the opportunity to see the patients, and were instructed to ask targeted rather than general questions (e.g. “do you have fever?” rather than “what other symptoms do you have?”). Therefore, it was expected that their
verbalizations reflected a relatively analytical approach to diagnosis. For this reason, additional data, reflecting a more intuitive diagnostic approach, were analysed for knowledge elicitation purposes. These were 24 protocols from a study where 18 family physicians were interviewed about patient cases that they believed to have diagnosed by intuition (Woolley and Kostopoulou 2013). At the interviews, the researchers prompted the physicians systematically, following the Critical Decision Method (CDM). The CDM has been used in numerous domains to investigate the cognitive components of proficient performance (Klein, Calderwood et al. 1989). It is a semi-structured interview method used to elicit information and knowledge from experienced users in relation to their decision making during non-routine, critical incidents. (Crandall, Klein et al. 2006) Using the CDM, the researchers elicited from the physicians the cues, expectancies, and goals associated with each judgment point. During the interviews, they also asked the physicians to identify potential errors at each decision point, and how and why errors might occur.

We used the situation assessment record (SAR) method to analyse both the think aloud and the interview protocols, to enable comparisons between them (Hoffman, Crandall et al. 1998). In SAR, the timeline for an event specifies the points at which the expert engaged in situation assessment and decision making. For each patient case, we constructed a chronological chart that showed how situation awareness evolved during the event: the types of knowledge, cues, interpretations and inferences that led to the situation awareness, and how situation awareness led to the course of action. Two examples of such a chart are presented in Table 1 and Table 2: table 1 depicts an excerpt from the SAR analysis of a think aloud protocol – the scenario features the first consultation of a patient with early myeloma; table 2 depicts an excerpt from the SAR analysis of a CDM interview – the scenario features a patient with ovarian cancer.
Eliciting user decision requirements for designing computerized diagnostic support for family physicians

*Analysis and representation:* Analysis and representation uses the results from the analyses of the previous phases and sets them out in a decision requirements table (Table 3), as suggested by the DCD framework.

**FINDINGS**

On the basis of the observations and follow-up interviews with 8 family physicians, we refined and expanded the diagnostic component of the HTA (Figure 1). Based on the HTA, intensive interviews, and analyses of the 34 think-aloud and 24 CDM protocols, we elicited four main cognitive requirements, which we encountered in both types of verbal protocols: 1) retrieving information from the patient record, 2) generating diagnostic hypotheses, 3) testing diagnostic hypotheses, 4) deciding on a patient management plan (Table 3). In most of the protocols, the initial situation assessment depended on retrieving information from the patient record and integrating it with the patient’s current health complaint (e.g., Table 1 – situation assessment 1). Throughout the diagnostic process, situation assessment depended on the generation and testing of diagnostic hypotheses (e.g., Table 1 – situation assessment 2 and 3; Table 2 – situation assessment 2). Physicians generate and test their hypotheses by asking the patient questions, performing examinations and ordering investigations, while constantly integrating and interpreting the information thus elicited.

Each one of the above tasks is also a cognitive requirement: physicians need to decide what information to elicit and when to stop eliciting more information and decide on a course of action, i.e., a “management plan”. Deciding on a management plan usually occurs towards the end of the diagnostic process (See Table 1 – situation assessment 3 and 4; Table 2 – situation assessment 3).

For each of the main four cognitive requirements, we reviewed all the diagnostic events to identify how the decision makers used cues, made inferences, and employed strategies to
fulfil the requirements. For each requirement, we then made interface design recommendations.

The different types of verbal protocols (think aloud and CDM), reflecting the intuitive-to-analytical spectrum of diagnostic reasoning, revealed the same key cognitive requirements in the diagnostic process. However, we identified different types of errors and strategies in the different types of protocol. For example, in the think-aloud protocols, we frequently identified omissions in information search (physicians not asking diagnostic questions/not performing important examinations or investigations). In the CDM protocols, on the other hand, we identified “sticking” to an initial diagnostic hypothesis as the most frequent cause of errors. Such hypotheses were based on, e.g., a colleague’s opinion, an earlier diagnosis, or previous knowledge about the patient leading to erroneous inferences about their current problem.

The process of transforming decision requirements into design recommendations is a critical one (Klein, Kaempf et al. 1997). For each requirement, we reviewed the diagnostic events to identify Human-Computer Interaction (HCI) concepts that could have provided useful support to the physician. We worked from the decision requirements themselves, but also went back to the protocols, to identify the type of information or perspective that could have made it easier to fulfil the requirement. This process enabled us to recommend HCI features for each decision requirement (Table 3). We limited the scope of the design recommendations to the diagnostic tool and its interaction with the EHR, and did not attempt to redesign the whole EHR system.

1. Retrieving information from the patient record

Family physicians must retrieve and integrate information from the patient record to build an understanding of the patient’s condition. They scan the EHR and its summary screen looking for significant information, either by browsing in a structured way (problems, medications,
previous consultations) or by actively filtering information (e.g., display only high priority problems). Important information may however be missed, if it is not well-presented and emphasised in the record or due to time constraints and distractions: “I missed once a cancer case. It was a woman in her 50s coming with a headache, I missed the information that when she was young she had cancer, it was in the record but at the very bottom, I didn’t scroll down” (physician 3, post-observation interview).

**Design recommendation**

To help retrieve and integrate critical information from the EHR that is relevant to the presenting problem, information should be displayed effectively, e.g., as text or icons, in the diagnostic support tool. Such critical information includes risk factors (smoking, excessive alcohol intake, hypertension) and serious past conditions, such as cancers, which are relevant to the patient’s current presenting problem. By making highly visible important patient information, the physician’s situation awareness can be supported better (Stanton, Chambers et al. 2001).

2. Generating diagnostic hypotheses

Physicians generated one or a small number of diagnostic hypotheses early in the consultation. “So it may be that she needs to be encouraged to be a little bit more patient or I’m thinking about disc prolapse” (Physician 9, think-aloud protocols, early myeloma scenario). Different factors, such as a colleague’s opinion, and assumptions about the patient (e.g., frequent consulter) may install a leading hypothesis at the exclusion of other alternatives. “I think I was agreeing with the earlier doctor, who saw her a week earlier, that maybe it was a sprain” ... “I think it was that feeling of...she comes here often, and she’s quite anxious because her husband left her recently and she was all alone and she’s struggling. And she wants reassurance that everything is doing okay” (Physician 1, CDM protocols, missed foot fracture).
Design recommendation

Displaying a list of potential diagnoses by integrating important information about the patient (e.g., age, gender, risk factors) from the EHR with the current health complaint could help physicians generate more diagnostic hypotheses. This could reduce narrow focus on one diagnosis developing early in the clinical encounter, expand the hypothesis space and remind physicians of other possibilities that should be considered. Following the results of two recent experimental studies where diagnostic accuracy improved over control with the mere presentation of a list of diagnostic hypotheses at the start of the consultation (Kostopoulou, Rosen et al. 2015) (Kostopoulou, Lionis et al.), a diagnostic support system could display a list of possible diagnoses as soon as the physician enters the patient’s main health complaint. The list could accommodate information entered before the consultation by other healthcare staff, such as physician assistants, or patients themselves.

3. Testing diagnostic hypotheses

In addition to deciding what questions to ask, examinations to perform and investigations to order, physicians have to decide when to stop gathering information and proceed to diagnosis and/or management. Physicians seeing the same patient can differ greatly in their diagnostic approach, as illustrated in the following example from the think-aloud protocols.

At the first patient visit, physician 3 asked 18 questions, performed 4 examinations and ordered 8 investigations before deciding to refer the patient to hospital: “So it’s becoming a bit more, looking like this lady may unfortunately have myeloma, which would fit with this persistent worsening back pain, mild anaemia and raised globulins and urine proteins. … So I’m referring her to the haematologist and I’m going to ask as an urgent…”

Physician 9 seeing the same patient asked only 3 questions and told her to come back if the back pain persisted. At the second patient presentation (with prolonged and worsening symptoms), physician 9 ordered a single investigation (X-ray of the back) and upon discovering that it was normal, the physician decided to prescribe pain relief: “Am I
concerned that there’s something that we’re missing or should I just try her for a bit longer with better analgesia? So I think, given that we haven’t tried better analgesia, I think that’s the next thing that I would do. So I’d stick with my diagnosis for now and increase her analgesia.”

This example illustrates two factors in the diagnostic process that may lead to error: first, that asking too few questions (presumably driven by a single hypothesis) may lead to misdiagnosis, as important information will not be discovered. Second, that once the physician adopts an interpretation, it may prove resistant to change, despite discovering new information that is inconsistent with that interpretation (Kostopoulou, Devereaux-Walsh et al. 2009) (Kostopoulou, Mousoulis et al. 2009). Information that is unexpected and/or cannot be easily integrated with the physician’s leading hypothesis may be dismissed or normalised: “His hemoglobin is absolutely fine” declared physician 1 while thinking aloud, after he/she was presented with an out-of-range hemoglobin, even though the abnormal result was marked with a * and the normal range was also provided next to it (“Hb 13.0 g/dL* - normal range 13.5-18 g/dL”).

**Design recommendation**

In addition to presenting a list of diagnostic suggestions, a support tool should enable users to click on a suggested diagnosis and view the important features (symptoms and signs) that can change the likelihood of the diagnosis. Users can check for these features in the patient and tick either “Yes” or “No” to indicate their presence or absence. The EHR will be updated automatically and so will the list of suggested diagnoses, if appropriate. For example, the order of the diagnoses may change according to their updated likelihood, and diagnoses may be added or removed. The tool should also propose examinations and investigations that could differentiate between the suggested diagnoses. In this way, physicians are likely to elicit and consider more information.
Data visualization based on the principles of Gestalt theory (gaining information “at a glance”) can support perception and situation awareness (Kim and Hoffmann 2003). Where appropriate, the system could contextualize abnormal or borderline investigation results according to patient demographics, risk factors and main health complaint, and present information in a combined visual display.

Coding information into the EHR is necessary for the operation of a diagnostic support system. If information is not coded or is entered in free text, the system cannot use it to support the diagnostic process in any interactive way, e.g. by updating its diagnostic suggestions. A diagnostic support system should therefore provide an easy interface for the coding of clinical information (symptoms and signs) during the consultation. In our observations of physicians consulting with patients, we noted that physicians recorded information either during the consultation or after the patient had left the room. Physicians may not record during the consultation, so that they concentrate on their interaction with the patient. Entering information after the patient has left, however, can result in loss of information and omission errors. Furthermore, we noted that physicians often did not code information but entered it as free text, which cannot be clinically interpreted by a computer. This reflects Salisbury and colleagues’ finding that 81% of problems discussed in consultations were recorded as free text and only 37% were coded (Salisbury, Procter et al. 2013).

Enabling physicians to indicate quickly either the presence or the absence of important features for specific diagnoses, as recommended above, can facilitate and encourage coding. Several other ways are facilitated by EHR systems, for example, hiding clinical codes that are redundant and not in use; auto-complete; providing default values and supporting quick access to previously inserted information; allowing keyboard shortcuts and the use of abbreviations. The most effective encouragement for the physician to code is likely to be the automatic transferring of the coded information in the appropriate locations of
the patient’s EHR. Another solution that also merits consideration and further research is for patients to enter their health complaint and associated symptoms on the computer, in the physician’s waiting room, before the start of the consultation.

Facilitating coding is important both for the acceptance of the tool by physicians and the specificity of its advice, as the coded information would be used by the diagnostic tool to update its list of suggested diagnoses. Furthermore, detailed recording of coded symptoms, including the main health complaint, would enable additional diagnostic evidence to be gathered and subsequently analyzed, consistent with the concept of the “Learning Healthcare System” (Friedman, Wong et al. 2010).

4. Deciding on course of action

Errors in management decisions can stem directly from misdiagnosis. They can also occur if the physician does not “safety net” for serious possibilities and only manages for what he/she considers to be the most likely diagnosis: “Normally, there’s lots of things that I didn’t do. To come back if worse – that’s the usual safety net; to come back if worse, covering myself” (Physician 2, CDM protocols). Finally, errors may occur due to insufficient knowledge about the most appropriate way to manage a specific disease. “So I’m going to be referring her for an urgent rheumatology review” (Physician 4, think aloud protocols, suspected of myeloma but referred to a rheumatology rather than oncology).

Design recommendation

When the physician enters a diagnosis, he/she should be able to link directly to the relevant clinical guidelines and forms for referring to specialists, ordering investigations and/or prescribing. This could be done by context-dependent information tools, such as Infobuttons. Infobuttons can be incorporated into the diagnostic support tool and integrate data about the patient and the clinical context to provide immediate, point-of-care access to relevant knowledge resources (Cimino, Li et al. 2002). Clicking on the Infobutton, next to the selected
diagnosis, could display to the physician specific information about the management steps that she/he should take. Selecting a step could then display the relevant form (e.g., request form for referral or investigation). The Infobutton would be linked to the latest clinical guidelines to provide accurate information at the point of care, and prevent management errors. Many UK EHR systems already provide an Infobutton or equivalent functionality, so this does not need to be redesigned into the diagnostic tool.

**DISCUSSION**

Using multiple methods and data sources, we elicited cognitive requirements of the diagnostic task and made specific user interface design recommendations for a computerized diagnostic support tool that will integrate seamlessly with the patient's EHR and will be triggered upon entry of the patient’s current health complaint. The tool will suggest diagnoses for physicians to consider early in the process, so that a narrow focus on a single hypothesis is lessened. The tool should also facilitate data coding and insertion, so that physicians enter more coded information into the EHR. It should suggest symptoms and signs that are important for the relevant diagnoses, and highlight significant information in the EHR. These features should enhance the tool’s usefulness and acceptability.

Existing design recommendations for decision support systems emphasize the importance of integration with the EHR, the consideration of the physicians’ workflow (Musen, Middleton et al. 2014), and support of the physicians’ cognitive tasks (Patel and Kaufman 2014). According to Stead and Lin’s (2009) National Academy of Sciences seminal report, current systems provide little support for the cognitive tasks and workflows of clinicians (Stead and Lin 2009). One of the report’s main conclusions is the need to provide “patient-centered cognitive support” that helps to “integrate patient-specific data where possible and account for any uncertainties that remain” (p. S-4). By eliciting and analysing family physicians’ decision making and cognitive requirements during the diagnostic process, our design recommendations ensure that cognitive challenges are addressed for the spectrum of
diagnostic reasoning, from the more intuitive to the more analytical. By integrating diagnostic support with the EHR and using patient-specific information to produce diagnostic recommendations, we are making an important step in the design of patient-centered cognitive diagnosis support.

Strengths and limitations

The strength of our work resides in its use of multiple and unique combination of methods for data collection and analysis (observations, interviews, HTA, CTA, and SAR) that helped to elicit cognitive user requirements of the diagnostic task. There are many CTA methods and tools available. Using a CTA entails selecting and applying a combination of methods and tools appropriate to the task and domain being investigated (Baxter, Monk et al. 2005). Combining different techniques is encouraged for eliciting requirements in software engineering (Nuseibeh and Easterbrook 2000). Nevertheless, most studies still tend to use one or a subset of these methods, while traditional techniques such as questionnaires and interviews are still most commonly used (Zowghi and Coulin 2005). CTA extends those traditional task analysis techniques to facilitate the collection of information about the cognitive processes underlying observable task performance (Chipman, Schraagen et al. 2000).

We used two different types of secondary data: think aloud protocols of family physicians diagnosing challenging cases (early presentations of cancer), and interview transcripts where family physicians described past cases of intuitive diagnoses. By using different types of secondary data, we wanted to ensure that the requirements elicited and design recommendations proposed are relevant to and can support the different modes of clinical thinking, from intuition to analysis (Evans and Stanovich 2013), on a range of clinical cases. For most of these cases, the correct diagnosis was known. It was therefore possible to identify errors and difficulties in the diagnostic process, e.g., important information omitted, hypotheses considered or not considered, and misinterpretations, which would have been
less likely by simply observing patient consultations in real time and relying on physicians’ self reports. This use of multiple data sources is novel in the requirements elicitation literature.

A limitation of our work is its focus on the diagnostic task at the exclusion of other tasks that family physicians routinely perform, e.g., prescribing. Future work to develop a fully functional diagnostic support tool will need to take into account how the EHR is used to manage other tasks during the consultation, and whether these tasks might be affected by the diagnostic tool. Furthermore, for our design recommendations to be effective, some physicians will need to change the way that they are currently interacting with their computer. For example, they will need to enter the patient’s current health complaint at the start of the consultation and read the diagnoses initially suggested by the system. The requirement for such behavioral changes is likely to increase resistance to system adoption. We believe, however, that a fully integrated and fully functional tool will offer substantial benefits to users, so that resistance is reduced and adoption motivation increased. Apart from the current benefits (integration with the EHR, ease of coding, automatic transfer of coded information into the EHR), a fully functional diagnostic support system would be driven by the latest diagnostic information, constantly updated, and be configured to allow patients and/or other caregivers to enter symptoms pre-consultation, so that time for the routine aspects of information gathering is reduced. In addition to these benefits, introduction of the system would require a carefully thought out “change management” plan to include time for training users on the system. Clear prerequisites for these next steps are that 1) an improvement in diagnostic accuracy is obtained first in a controlled environment (a relevant study is underway), 2) the evidence base driving the tool (symptoms/signs and their link to diagnoses) is rich and trusted by physicians, and 3) the tool’s usability is improved, for example, by less cumbersome ways of data entry.
We envisage (and propose the first steps in) the development of a “Learning Health System” for diagnosis, where a cycle of evidence-based quality improvement is created, use of the tool supports better coding and structure of routine diagnostic data, the data are made available to researchers to analyse and enrich the clinical diagnostic evidence, and the evidence is fed back into the tools’ recommendations to support better decision making.

**Funding:** This study received funding from the European Union’s Seventh Framework Programme for research, technological development and demonstration under grant agreement no 247787 [TRANSFoRm]. Financial support for the study that elicited the think-aloud protocols was provided by Cancer Research UK to Olga Kostopoulou, under the NAEDI scheme (ref. C33754/A12222). Financial support for the study that elicited the CDM interview protocols was provided by a departmental PhD studentship to Amanda Woolley.

**Ethical approval:** Ethical approvals were obtained from the North West London Research Ethics Committee 2 (10/H0720/50) and West London Research Ethics Committee 2 (11/LO/0079).

**Acknowledgements:** Drs Ellen Wright and Thomas Round, family physicians and clinical academic fellows at King’s College London, provided clinical advice and helped with piloting the follow up interviews and recruiting physicians for observation.
References

Eliciting user decision requirements for designing computerized diagnostic support for family physicians


Eliciting user decision requirements for designing computerized diagnostic support for family physicians


Table 1: Excerpt from the SAR analysis of a think-aloud protocol of physician 9
(patient with early myeloma)

<table>
<thead>
<tr>
<th>Situation Assessment 1</th>
<th>Cues</th>
<th>Expectations</th>
<th>Goal</th>
<th>Decision point 1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>67, female, a bit on the heavy side, has a history of blood pressure, hypertension and arthritis. She’s on medication for her blood pressure and she’s been seen with a back problem relatively recently. Referred for physio and note that she attends quite frequently. Seems to be well but is holding her back and does seem to be in pain. So she’s talking about having back pain. She’s taken cocodemol and the pain hasn’t got better.</td>
<td>I’m assuming that she’s wanting better pain relief. She is already taking co-codamol, she’s had the physio.</td>
<td>Ask about the physio / explore the reason for encounter</td>
<td>“Did you feel that the physiotherapy made any difference or whether the treatment is still ongoing?”</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Situation Assessment 2</th>
<th>Cues</th>
<th>Hypotheses 1 and 2</th>
<th>Goal</th>
<th>Decision point 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>“I’ve seen the physio twice now. She recommended some exercises to do at home. I try to do them every day. They keep me active but I am not sure how much they are helping. I have some follow up sessions booked.”</td>
<td>“So it may be that she needs to be encouraged to be a little bit more patient or I’m thinking about disc prolapse”</td>
<td>Differentiate between hypotheses 1 and 2</td>
<td>“Has the pain got any worse?”</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Situation Assessment 3</th>
<th>Cues</th>
<th>Expectations</th>
<th>Hypothesis 2</th>
<th>Goal</th>
<th>Decision point 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>“After the pain started it has been the same constant aching most days.”</td>
<td>“I think she’s only had two sessions of the physio, she’s only been doing the exercises at home for a little while. I’m thinking she probably just needs to be a bit more patient”.</td>
<td>Mechanical pain (“So it does sound like it’s mechanical low back pain”).</td>
<td>Increase medication to control the pain</td>
<td>“Ask her how much of the co-codemol she’s taking and if she’s getting any problems with it such as constipation”.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Situation Assessment 4</th>
<th>Cues</th>
<th>Hypothesis 2</th>
<th>Goal</th>
<th>Decision point 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>“I have been taking the co-codamol tablets, as I was told, mostly 8 per day. They take the edge off the pain but it doesn’t go completely”.</td>
<td>Mechanical low back pain</td>
<td>“She’s taking up to the maximum dose so I don’t particularly want to change anything there”.</td>
<td>“She should continue with her exercises and give the physio a bit more chance, so I’ll ask her to come back if it doesn’t settle and to continue with her physio”.</td>
</tr>
</tbody>
</table>
Table 2: Excerpt from the SAR analysis of a CDM protocol of physician 8 (patient with ovarian cancer)

<table>
<thead>
<tr>
<th>Situation Assessment 1</th>
<th>67, female, not happy with previous consultation, felt unwell, tired with no energy, dizzy, loss of appetite, weight loss (a stone over a year), overweight, ex-smoker, previous tests were all normal. Non-diagnostic cues: “A very, very frequent attender… and has multiple social and psychological problems. And she wears you out…”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expectations</td>
<td>She lost a stone in weight over the last year, less concerning since the weight loss was over a long period of time (“but if you lost that sort of weight over a month or two or three, then you’re going to be worried”), in addition, “she is a biggish lady, and therefore a stone didn’t feel that much in that sense. But she’d never lost weight before. And not many 67 year olds go on a diet. I mean, young women and men do, but not a 67 year old”.</td>
</tr>
<tr>
<td>Decision point 1</td>
<td>Perform a full examination</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Situation Assessment 2</th>
<th>Abdomen normal, chest normal (full examination normal)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expectations</td>
<td>“I think my guts told me that there wasn’t anything too much wrong with this lady”</td>
</tr>
<tr>
<td>Hypothesis 1</td>
<td>Seemingly nothing is wrong, possibly irritated bowel</td>
</tr>
<tr>
<td>Goal</td>
<td>Confirm nothing is wrong with blood tests</td>
</tr>
<tr>
<td>Decision point 2</td>
<td>Perform screen tests (to look at all options): blood tests (ESR, kidney and liver, sugar, thyroid, FBC) and chest x-ray.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Situation Assessment 3</th>
<th>Blood tests normal, chest x-ray normal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expectations</td>
<td>Given problem with previous consultation best to follow through to avoid upset: “I did pick up a feeling in her voice that she was unhappy with the previous consultation with my partners, and I thought by following her through and making sure that nothing went wrong, as it were, that we might avoid any complaint or upset. So that was really why I brought her back”.</td>
</tr>
<tr>
<td>Hypothesis 2</td>
<td>Nothing serious</td>
</tr>
<tr>
<td>Goal</td>
<td>Arrange a follow-up</td>
</tr>
<tr>
<td>Decision point 3</td>
<td>Arrange an appointment in two weeks</td>
</tr>
</tbody>
</table>
### Table 3 – Decision Requirements Table

<table>
<thead>
<tr>
<th>Decision/cognitive requirement</th>
<th>Difficulty/potential errors</th>
<th>How is it done?</th>
<th>Critical cues</th>
<th>Design recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Retrieving information from the Electronic Health Record (EHR)</td>
<td>Important information in the EHR may be missed.</td>
<td>Physicians scan the EHR looking for significant information They retrieve, assess and integrate the information.</td>
<td>Risk factors, chronic diseases and serious past conditions relevant to the current presenting problem.</td>
<td>Make critical information more salient in the diagnostic tool.</td>
</tr>
<tr>
<td>2. Generating diagnostic hypotheses</td>
<td>Focusing narrowly on a hypothesis, at the exclusion of other alternatives.</td>
<td>Physicians generate a small number of diagnostic hypotheses early in the consultation.</td>
<td>As above. Also, situational factors, e.g. patient appearance or behaviour, frequency of previous consultations.</td>
<td>Display a list of potential diagnoses by integrating information from the EHR with the current health complaint. The list will be displayed as soon as the physician types a complaint into the EHR.</td>
</tr>
<tr>
<td>3. Testing diagnostic hypotheses</td>
<td>Difficulty integrating new information that is not consistent with the leading hypothesis.</td>
<td></td>
<td>Diagnostic hypotheses, patient input, history taking, physical examination, current and past investigations.</td>
<td>Update the list of diagnoses according to coded input of clinical information. Provide interface for entering coded information quickly and easily.</td>
</tr>
<tr>
<td>3.1. Deciding what questions to ask the patient</td>
<td>Not asking diagnostic questions. Asking too few questions.</td>
<td>Physicians ask about “alarm” symptoms to rule out serious diagnoses.</td>
<td>Diagnostic hypotheses and patient symptoms.</td>
<td>Enable physicians to click on a suggested diagnosis and view its features that can alter diagnostic likelihood.</td>
</tr>
<tr>
<td>3.2. Deciding if and what examinations to perform/investigations to order</td>
<td>Not examining or investigating appropriately (ordering too few or too many investigations).</td>
<td>Physicians request investigations to rule out serious disease but also in case of uncertainty.</td>
<td>Diagnostic hypotheses, patient symptoms.</td>
<td>Propose examinations and investigations to differentiate between the possible diagnoses.</td>
</tr>
<tr>
<td>3.3. Interpreting the information elicited</td>
<td>Normalising or dismissing symptoms that do not fit with the leading hypothesis.</td>
<td>Physicians interpret information according to their leading hypothesis.</td>
<td>Diagnostic hypotheses, and the diagnostic value of information.</td>
<td>Contextualise abnormal or borderline results according to patient’s age, main health complaint and possible diagnoses. Highlight recurrent symptoms and relevant, past test results. Data visualisation.</td>
</tr>
<tr>
<td>3.4. Deciding when</td>
<td>Asking too few</td>
<td>Stop searching,</td>
<td>Diagnostic</td>
<td>Display diagnostic</td>
</tr>
<tr>
<td>Decision/cognitive requirement</td>
<td>Difficulty/potential errors</td>
<td>How is it done?</td>
<td>Critical cues</td>
<td>Design recommendations</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>----------------------------</td>
<td>----------------</td>
<td>---------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>to stop eliciting information</td>
<td>questions. Omitting diagnostic questions.</td>
<td>once a satisfactory explanation is reached and the most serious alternatives have been explored.</td>
<td>hypotheses.</td>
<td>cues for each suggested diagnosis to remind physicians of the key questions that they need to ask.</td>
</tr>
<tr>
<td>4. Deciding on course of action</td>
<td>Inappropriate management due to misdiagnosis or insufficient knowledge about how best to manage a specific condition.</td>
<td>Physicians follow guidelines where available.</td>
<td>Latest clinical evidence and guidelines. Patient demographics and risk factors.</td>
<td>After physicians enter a diagnosis, an Infobutton can link to the most appropriate next steps, such as forms for referral, investigations and prescribing, in relation to that diagnosis.</td>
</tr>
</tbody>
</table>
Figure 1: The DCD framework – sources of data, methods of analysis, and outputs of each DCD phase.
0 Diagnosing a patient

PLAN 0: before starting surgery do 1, 2, 3 and 4 in any order and combination. If patient has arrived and when ready to receive patient - 5. When patient comes into the consultation room - 6, then 7 and 8. If do not know or cannot remember patient’s information go back to 4. Throughout the consultation do 11. If decided on course of action do 9 and 10. If further information regarding present visit needs to be obtained – 7. If next patient arrived and when ready to receive patient – 5 and carry on as for previous patient.

1. Review the list of the day’s appointments
2. Check whether first patient has arrived (system indication)
3. Open patient health record
4. Get familiar with patient’s clinical history
5. Call patient in
6. Establish why patient has come
7. Obtain new information

PLAN 7:

Do 7.1. As appropriate and if available – 7.2 and 7.3, in any order and combination.

7.1 Obtain new information from patient
- INTERPRETING THE INFORMATION ELICITED
- DECIDING WHEN TO STOP ELICITING INFORMATION

PLAN 7.1

Do any 7.1.1 to 7.1.4 in any order and combination, as appropriate.

7.1.1. Obtain information verbally
PLAN 7.1.1

Do 7.1.1.1 and 7.1.1.2 thoroughly

7.1.1.1. Ask the patient questions
- DECIDING WHAT QUESTIONS TO ASK
7.1.1.2. Listen to patient’s answers

7.1.2. Obtain information via observation
PLAN 7.1.2

Do 7.1.2.1 and 7.1.2.2 in any order and combination.

7.1.2.1. Observe patient’s appearance (e.g., skin colour)
7.1.2.2. Observe patient’s behaviour (e.g., apathy in children).

7.1.3. Obtain information via physical examination
- DECIDING IF AND WHAT EXAMINATIONS TO PERFORM
PLAN 7.1.3

Do any 7.1.3.1 to 7.1.3.2 in any order and combination, as appropriate.

7.1.3.1. Perform general checks (e.g., blood pressure, pulse).
7.1.3.2. Perform specific checks according to the current health complaint (e.g., abdominal examination, chest examination).

7.1.4. Obtain information via near-patient testing (e.g., dipstick urine analysis)
7.2 Review relevant pathology results
7.3 Review relevant clinical correspondence
8. Decide on course of action
9. Provide “Safety net” (e.g., ask patient to come back if condition does not improve)
10. Carry out next steps
11. Record consultation