

Translational Research and Patient Safety in Europe

The TRANSFoRm Project is partially funded by the European Commission under the 7th Framework Programme

Grant Agreement Number FP7-247787

D2.1

Most promising type of support for diagnosis

Work Package Number:	2
Work Package Title:	Decision Support use case
Nature of Deliverable:	Report
Dissemination Level:	Public
Version:	1
Delivery Date From Annex 1:	M36
Principal Authors:	Olga Kostopoulou, Brendan Delaney
Contributing Authors:	Andrea Rosen, Thomas Round, Ellen Wright
Partner Institutions:	King's College London (KCL) Royal College of Surgeons in Ireland (RCSI) University of Crete (UoC)

7th Framework Programme <http://cordis.europa.eu/fp7/ict/>

European Commission http://ec.europa.eu/information_society/

Table of Contents

Table of Contents	2
List of Figures	3
List of Tables.....	3
Executive Summary.....	4
1. Introduction	7
2. Methods	9
2.1 Materials	9
2.2 Simulating diagnostic support.....	10
2.3 Sample size	11
2.4 Participants	11
2.5 Procedure	12
2.6 The study in Greece.....	16
3. Results	18
4. Discussion	21
5. Conclusions.....	23
6. Next steps	23
References	24
Abbreviations	26
Appendix	27

List of Figures

Figure 1: The initial screen that all GPs saw: example of a scenario. 13

Figure 2: The answer to a GP’s question is presented at the bottom of the screen, below the patient information (example of a scenario). 14

Figure 3: The list of suggestions seen by the early support group (example of a scenario). 15

Figure 4: The list of suggestions seen by a participant in the "alerting" group, once he/she entered a diagnosis and management for one of the scenarios. 16

List of Tables

Table 1: The correct diagnosis (underlined) and its main diagnostic competitor in each scenario. 9

Table 2: Mean (SD) of correct diagnoses out of 9 per experimental condition for the UK and Greek samples. 18

Table 3: Means (SD) of information gathered by experimental condition for UK and Greek GPs. 20

Executive Summary

Work Package 2 (“Decision Support use case”) aimed to investigate the most appropriate timing and type of diagnostic support for the primary care setting and elicit user requirements to be used in the subsequent development of a prototype diagnostic support system (DSS), integrated with the electronic health record (EHR). The first work task (*WT 2.1 - Development of materials for studying diagnosis*) involved the development of materials for an experimental study that compared two different principles of diagnostic support (early support and late support) against a control group that received no support (*WT 2.2 - Comparison of two different generic approaches to diagnostic support*). Nine evidence-based and rich-in-detail clinical scenarios were created to cover a range of diagnostic difficulty. Scenario patients presented with one of three reasons for encounter: chest pain, abdominal pain or dyspnoea. The clinical content of the scenarios was supported by literature review, carried out by the Royal College of Surgeons in Ireland (RSCI) as part of WT 4.1 (*Repository of clinical prediction rules*), as well as expert opinion. The results of this Deliverable are being used in WT 2.3 (*User requirements specification*) and WT 5.2b (*Specification for a functional DSS data collection tool to be integrated with an EHR*) to guide the design of the DSS prototype, which will be evaluated in WT 2.5 (*Validation of the final DSS prototype in realistic clinical situations*).

A computerised interface was designed by King’s College London (KCL) to present the scenarios to GPs and collect the data remotely over the Internet using a web tool. Participating GPs could gather information about the simulated patients by asking questions that a researcher answered (i.e., the researcher selected the appropriate answer from a list of predetermined cues and sent it to the participant’s screen in real time). Participants and researcher were in concurrent phone communication. The web tool automatically recorded all questions participants asked in sequence and the time taken.

Early diagnostic support involved presenting participants with a list of diagnostic suggestions as soon as they read some background patient information (age, sex, risk factors, current medications, past medical history) and the reason for encounter. The list remained on screen for a minimum of 20 seconds and disappeared as soon as the participants started their information gathering in order to diagnose and

manage the scenario patients. This early support aimed to help participants consider more diagnostic hypotheses than they otherwise would and reduce the tendency quickly to narrow down on one or a small number of hypotheses. Participants could make the list of diagnostic suggestions appear again on their screen at any point they wished during the consultation with a patient.

Another group of participants received a list of diagnostic suggestions only after they had gathered information and had given their own diagnosis and management. This list depended on the information each participant had gathered during the consultation. It aimed to alert participants to diagnoses that they had not excluded and represented the more traditional approach to diagnostic support, where physicians first gather information as they see fit and then enter all the information into the system to receive advice about differential diagnoses. After seeing the list, participants could request further information, change their diagnosis or give the same diagnosis.

The lists of diagnostic suggestions for each scenario were derived from an existing, stand-alone, diagnostic support system for internal medicine, DxPlain (<http://dxplain.org/>), and were adapted to primary care by the KCL team. These lists constituted the early support lists for each scenario. The late support lists originated from the early support lists, which were narrowed down depending on the information each participant had gathered. The KCL team formulated rules that determined which diagnoses from the early support lists could reasonably be discounted, had a participant requested specific cues. Thus, late support was individualised to each participants' information search.

The study initially took place in the UK. All study materials were translated and adapted to a Greek context by the teams at the University of Crete (UoC) and KCL. The study was repeated in Greece with a smaller sample of GPs in order to ascertain the feasibility and likely effectiveness of diagnostic support in a country with a newer and substantially different primary care system than the UK. There was significant heterogeneity between the studies in the two countries, therefore, the data were not pooled but were analysed separately.

297 GPs in the UK and 150 GPs in Greece were allocated via blocked randomisation

to one of three experimental conditions: control, early support or late support. The primary outcome was diagnostic accuracy. In the UK, early support significantly improved diagnostic accuracy over control (OR 1.31 [95% CI 1.04-1.66], $P=0.023$), while late support did not (OR 1.10 [95% CI 0.87-1.38]). In Greece, both types of support significantly improved accuracy over control: OR 1.68 [95% CI 1.23-2.30] ($P=0.001$) for early support and OR 1.50 [95% CI 1.10-2.04] ($P=0.011$) for late support.

Reminding GPs of diagnostic possibilities to consider early on in the consultation, before they start narrowing down on a hypothesis, can significantly improve diagnostic accuracy across a wide range of diagnostic scenarios, without significantly increasing the amount of information gathered. In a sample of GPs in Greece, both early and late reminders improved diagnostic accuracy over control. Further research in Work Packages 2, 4 and 5 will thus focus on developing a DSS prototype, integrated with the EHR, that has the potential to provide early reminders in response to the reason for encounter but is also flexible enough to make diagnostic suggestions at other points during the consultation and with varying amounts of information.

1. Introduction

Diagnostic error is the commonest cause of litigation against GPs.^{1 2} Diagnostic delay can harm patients, especially in life-threatening or progressive conditions. Proposed solutions include training with clinical scenarios, formal teaching of clinical reasoning and computerised diagnostic support.³ Computerised systems for disease management, preventive care and prescribing are used extensively in clinical practice on both sides of the Atlantic.⁴ In contrast, diagnostic support systems (DSS) have not enjoyed similar success, despite more than four decades of development.⁵ Most studies have evaluated the performance of specific DSS, i.e., whether they can generate the correct diagnosis for challenging cases,^{6 7} rather than the performance of physicians using the systems, while randomized studies are not carried out.⁸⁻¹⁰

The basic operation of the typical DSS has remained the same throughout its history: the physician collects information about the patient, enters the information collected into the DSS and receives a list of diagnostic suggestions to consider. There are two problems with this approach. First, it requires that the physician decide to consult the system. Physicians, however, do not always know when advice would help and do not always recognize when a case is difficult.^{11 12} In a naturalistic trial of ISABEL, a physician-triggered DSS, junior doctors in paediatric ambulatory care sought and examined the system's advice only around 2% of the time.¹⁰ It remains uncertain whether automated provision of decision support is more effective than physician-triggered decision support, as comparisons of this feature have not included diagnostic systems. As a general rule, systems that simply generate alerts are likely to be overridden by physicians, especially when multiple alerts are built into an EHR system.^{4 13 14}

The second problem with the current approach to diagnostic support is that system advice comes late in the diagnostic process. Physicians generate a small number of diagnostic possibilities from the start of the encounter (within seconds) and these determine what information they will gather and how they will interpret it.¹⁵⁻¹⁸ Consequently, advice given late in the consultation, after a fair amount of information has been gathered, may be less effective in two ways. First, the information a physician will enter into the DSS and its resulting advice will depend on the hypotheses already entertained.¹⁹ Important information may be missed out or normalized.¹⁷ Second, once physicians have mentally represented the problem and considered a potential cause, a cognitive set may develop,²⁰ making them less open to the system's suggestions and less likely to change their mental representation of the patient's problem. Therefore, a potentially more successful approach would be to present diagnostic suggestions as early as possible in the consultation, before physicians have started testing any diagnostic hypotheses. Such early suggestions could be triggered automatically, based on the reason for encounter and information in the patient's record.

To test whether providing physicians with hypotheses early in the process improves diagnostic accuracy, we constructed detailed patient scenarios and presented them to GPs to diagnose and manage via a web tool. We compared the performance of GPs who received early diagnostic suggestions with that of an unaided group of GPs (control) and that of a group that received diagnostic suggestions late in the process, based on the information each GP had gathered. This represented the traditional approach to diagnostic support.

2. Methods

2.1 Materials

Chest pain, abdominal pain and dyspnoea are common reasons for consulting the GP. They can be caused by a multitude of conditions, some serious. Using a series of evidence-based reviews conducted by RCSI (WT 4.1), we developed nine patient scenarios, three for each reason for encounter (WT 2.1), ranging in difficulty. The scenarios depicted conditions that, although not common, had sufficiently high prevalence to be considered relevant to primary care. In each scenario, a more common diagnosis could explain some but not all of the patient’s symptoms and signs (Table 1).²¹ We expected participants to think of this diagnosis promptly and relatively easily. If participants missed the intended diagnosis, the patient could suffer serious clinical consequences. Each scenario contained background information about the patient, a reason for encounter and an exhaustive list of available positive and negative diagnostic information (see Appendix for an example). Scenarios were extensively piloted to test their plausibility and ensure that most GP queries for information could be answered.

	Easy	Moderately difficult	Very difficult
Chest pain	<u>Angina</u> vs. musculoskeletal pain	<u>Pulmonary embolism</u> vs. lower respiratory tract infection	<u>Tuberculosis</u> vs. lower respiratory tract infection
Abdominal pain	Crohn’s disease vs. enteritis	<u>Appendicitis</u> vs. urinary tract infection	<u>Ovarian cancer</u> vs. irritable bowel syndrome
Dyspnoea	Childhood asthma vs. bronchitis	<u>Cor pulmonale</u> vs. COPD exacerbation	COPD and aortic stenosis vs. COPD alone

Table 1: The correct diagnosis (underlined) and its main diagnostic competitor in each scenario.

2.2 Simulating diagnostic support

To determine the relevant diagnoses for each scenario and ensure completeness, we adapted diagnostic suggestions from DXplain, a stand-alone decision support system designed for general internal medicine (<http://dxplain.org>). We entered into DXplain the background information about each scenario patient (age, sex, risk factors, current medications, past medical history) and the reason for encounter (chest pain, abdominal pain or dyspnoea). DXplain then delivered a list of suggested diagnoses, which was scrutinized by two GPs to ensure its appropriateness for UK primary care. The average list length across scenarios was 17 diagnoses (range 9-22) and always contained the correct diagnosis (see example in Appendix). These diagnostic lists were presented to the early support group at the start of each scenario.

The late support group did not receive any support until after they made their diagnosis. They were then presented with a list of diagnoses to consider. These were diagnoses that could still not be discounted. These diagnoses originated from the respective early-support lists, after applying predetermined exclusion rules. We formulated these exclusion rules for each scenario via clinical consensus. The rules determined the diagnoses that could be reasonably discounted, if GPs had requested specific scenario cues. For example, for the patient that presented with a few weeks' chest pain, if a GP had checked for chest wall tenderness (response: *"No chest wall tenderness on examination. Pressure on chest wall does not reproduce pain."*), we assumed that costochondritis could be discounted. Thus, after a diagnosis was entered, each GP in the late support group received an individualised list of diagnostic suggestions, based on what information he/she had requested during the encounter.

2.3 Sample size

Sample size calculation was based on a previous study that used a similar data collection method: 84 GPs diagnosed 7 challenging scenarios without diagnostic support. Mean diagnostic accuracy was 0.42 (SD 0.22), representing the expected accuracy of the control group.²¹ We estimated a slightly increased standard deviation of 0.25 to account for 6 rather than 7 difficult scenarios. To detect a 10% increase in mean diagnostic accuracy with diagnostic support from baseline with a power of 80% and a significance level of 0.05 in a two-sample t-test, 99 participants per group would be needed.


2.4 Participants


Practices across England were invited to participate in the study via the NIHR Primary Care Research Network (http://www.crncc.nihr.ac.uk/about_us/pcrn). Their GPs could contact the study team, if they wished to participate. Participants were offered funding at standard GP rates for an estimated 3-hour involvement and individualised feedback, which they could use towards professional development requirements. We recruited 297 GPs, including 30 GP registrars (trainees in general practice), to reflect the proportion of registrars in the UK GP population. 54% were female and the average number of years in general practice post-qualification was 8.77 (SD 9.32, range 0-34), being lower than the UK average of 16 years. All UK GPs are familiar with using an EHR. All practices in the UK use them, some for more than 20 years. Experience of DSS is nevertheless absent.

2.5 Procedure

Participating GPs saw the 9 scenarios in random order, in one of three experimental conditions: control, early support or late support. Assignment to experimental conditions followed a pre-determined blocked randomisation sequence that ensured equal numbers of participants per condition.

Data collection took place remotely over the Internet using a web-tool designed specifically to present the scenarios and the diagnostic suggestions, and collect the data. Participants were in simultaneous phone communication with a researcher, who operated the site and guided them through the task during a single session. After receiving training on one scenario, participants proceeded to diagnose and manage the 9 scenarios. At the start of each scenario, they read the introductory information about the patient: name, age, sex, BMI, risk factors such as blood pressure, alcohol and smoking, current medications, co-morbidities, reason for the last recorded encounter, and a brief description of their appearance. On the same screen, they read the reason for encounter (Figure 1). After confirming that they had read this information, participants could request more information, in relation to history, physical examination and investigations. For each scenario, we had prepared a list of answers (“cues”) to potential questions. After each question, the researcher chose the appropriate cue and displayed it on the participant’s screen (Figure 2). When participants wished to finish the consultation, they entered the diagnosis that they considered most likely for the patient and selected their management decision from a list of options (refer, prescribe, arrange follow-up, advice or wait and see). The system automatically recorded all information requests in sequence, the timing of each request, the diagnoses and management decisions. No feedback was provided to participants until all data collection for the study was completed.





Mandy Smith

Patient information

- AGE: 28 years old
- ETHNICITY: Caucasian
- HEIGHT: 1.62 m
- WEIGHT: 55 kg (BMI 20, measured 12 months ago)
- ALCOHOL: 12 units per week
- SMOKING STATUS: Never smoked
- LAST BP: 120/80, taken 12 months ago
- PAST MEDICAL HISTORY: None
- MEDICATION: None
- LAST CONSULTATION: For abdominal pain, 4 months ago
- APPEARANCE: She looks a little unwell and pale as she comes in.

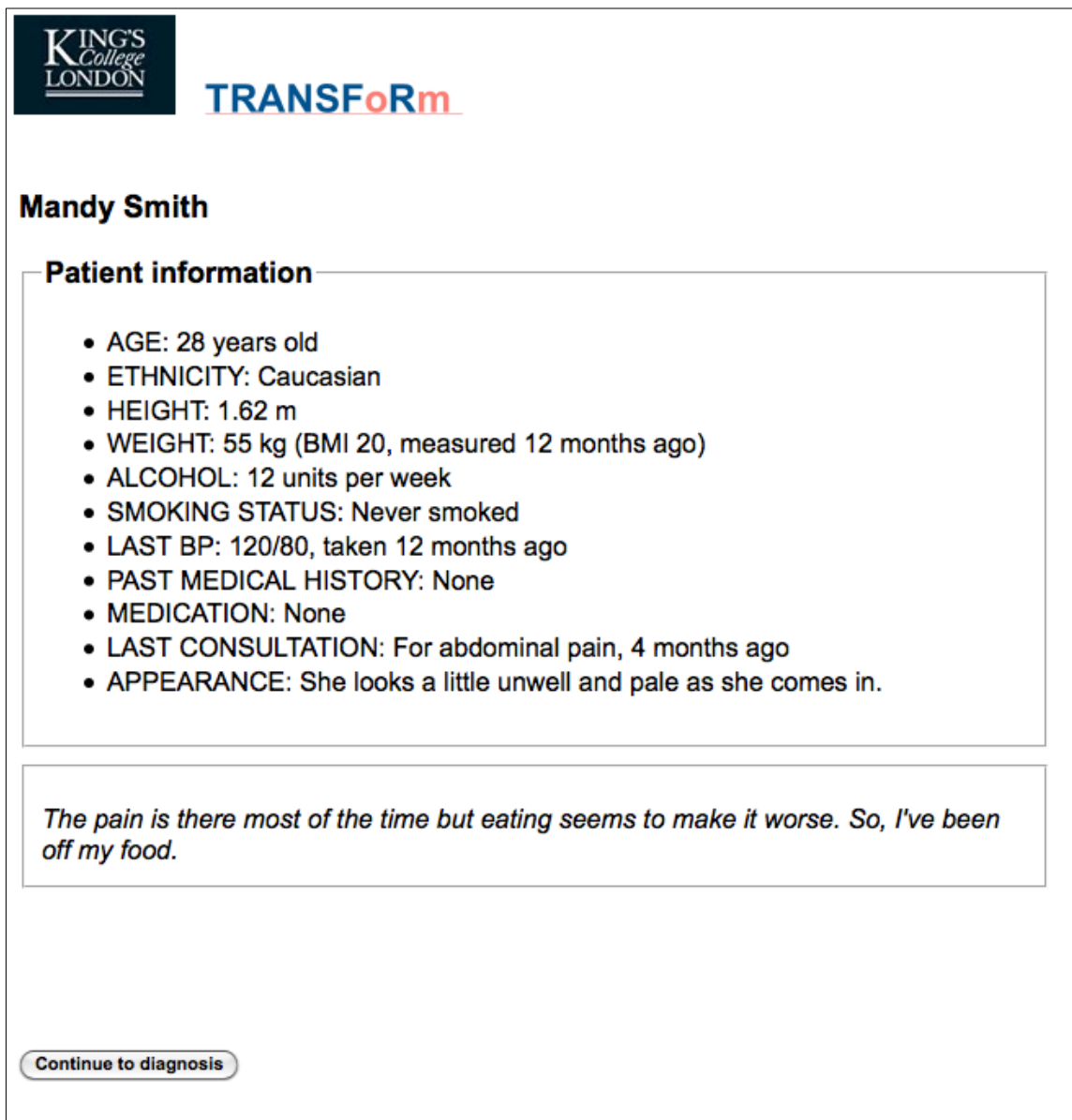
Presenting complaint

"Good morning doctor. How are you today? I haven't been well for about a week. I've been off work with a really bad stomach pain and diarrhoea. I thought that I had picked up a bug and that it would get better by itself but it hasn't."

Confirm you have read the Presenting Complaint

Waiting for researcher...

Figure 1: The initial screen that all GPs saw: example of a scenario.



KING'S
College
LONDON

TRANSFoRm

Mandy Smith

Patient information

- AGE: 28 years old
- ETHNICITY: Caucasian
- HEIGHT: 1.62 m
- WEIGHT: 55 kg (BMI 20, measured 12 months ago)
- ALCOHOL: 12 units per week
- SMOKING STATUS: Never smoked
- LAST BP: 120/80, taken 12 months ago
- PAST MEDICAL HISTORY: None
- MEDICATION: None
- LAST CONSULTATION: For abdominal pain, 4 months ago
- APPEARANCE: She looks a little unwell and pale as she comes in.

The pain is there most of the time but eating seems to make it worse. So, I've been off my food.

Continue to diagnosis

Figure 2: The answer to a GP's question is presented at the bottom of the screen, below the patient information (example of a scenario).

The procedure above outlines the process for the control group. The early support group followed the same process with one main difference. After participants confirmed that they had read the introductory information about the patient and the reason for encounter, they were presented with a list of diagnostic suggestions (Figure 3). The list remained on the screen for a minimum of 20 seconds. In order to proceed, participants had to confirm that they had read the list. This then

disappeared from the screen and they could start asking questions about the patient. They could make the list appear on their screen at will during the consultation.

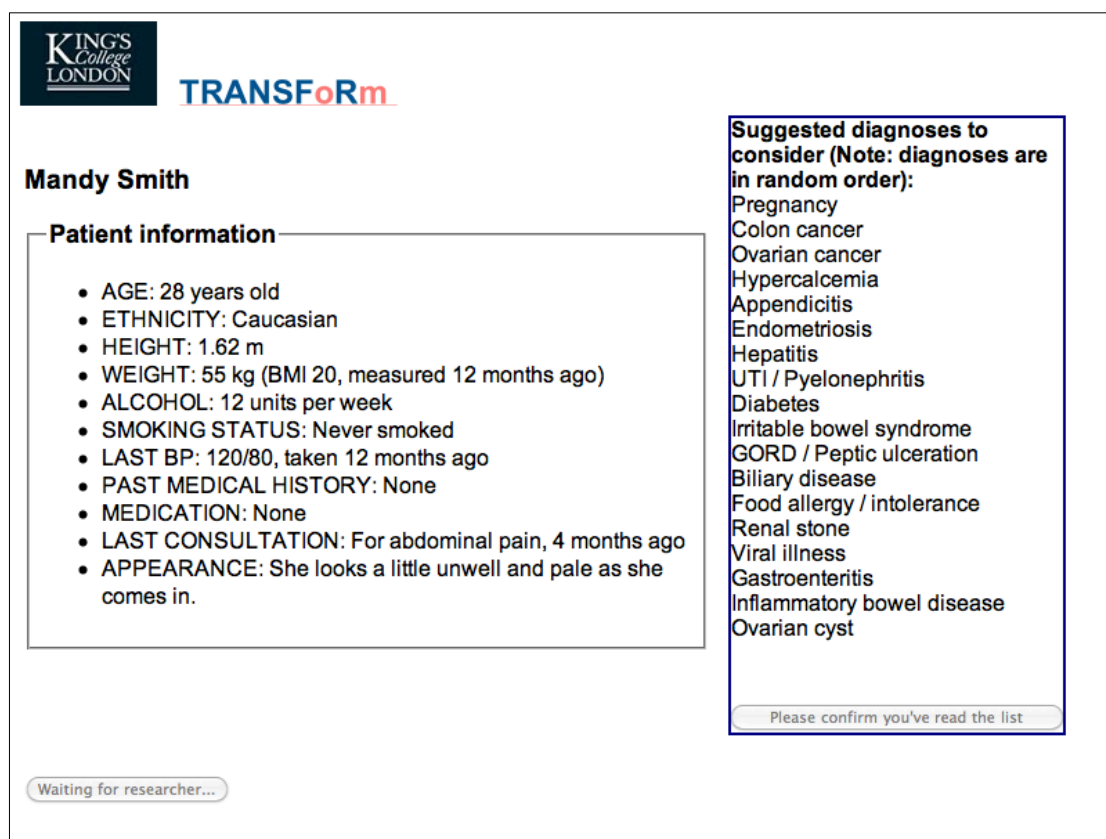
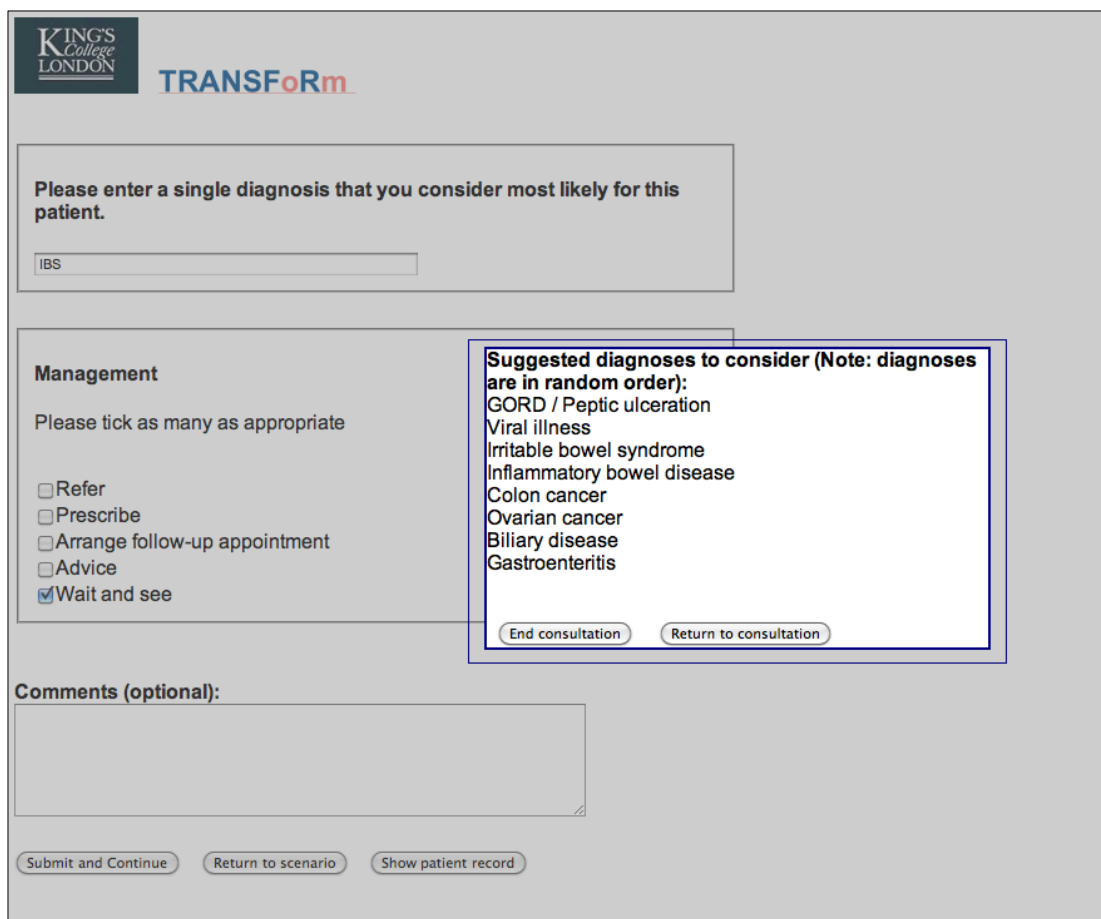


Figure 3: The list of suggestions seen by the early support group (example of a scenario).

The late support group followed the same process as the control group and submitted their diagnosis and management. This triggered the list of diagnostic suggestions (Figure 4). Physicians could then choose to ask more questions about the patient and/or change their diagnosis if they wished. The diagnostic suggestions in both early and late support were presented in random order for each participant.



KING'S
College
LONDON

TRANSFoRm

Please enter a single diagnosis that you consider most likely for this patient.

IBS

Management

Please tick as many as appropriate

Refer
 Prescribe
 Arrange follow-up appointment
 Advice
 Wait and see

Suggested diagnoses to consider (Note: diagnoses are in random order):
GORD / Peptic ulceration
Viral illness
Irritable bowel syndrome
Inflammatory bowel disease
Colon cancer
Ovarian cancer
Biliary disease
Gastroenteritis

End consultation Return to consultation

Comments (optional):

Submit and Continue Return to scenario Show patient record

Figure 4: The list of suggestions seen by a participant in the "alerting" group, once he/she entered a diagnosis and management for one of the scenarios.

2.6 The study in Greece

All study materials (invitation letters, scenarios, diagnostic suggestions per scenario and task instructions) were translated to Greek by the UoC and KCL research teams (WT 2.1). The study was performed in Greece with Greek GPs as participants. 150 GPs from all over Greece participated. Given the smaller size of Greek GP population in comparison to the UK, a third of the UK sample was considered feasible and adequate for assessing the feasibility and likely effectiveness of a DSS in Greece.

The sample of 150 Greek GPs included 30 GP registrars. 43% were female and the average number of years in general practice post-qualification was 5.34 (SD 5.35, range 0-23).

To score GPs' management decisions, we considered whether patient harm could result from delay in dealing with the condition in each scenario and predefined the type of decisions that would be considered appropriate or inappropriate in each scenario. The scheme for scoring the management decisions was also translated to Greek and the UoC team scored the individual management decisions of the 150 participants. The diagnoses given by the Greek GPs were translated into English by the UoC team and the KCL team scored them as either correct (if they matched the diagnoses depicted in the scenarios) or incorrect. The KCL team performed all the data analyses.

3. Results

Significant heterogeneity was found between the studies in the UK and Greece. Specifically, a fixed-effects meta-analysis of the relative risk of diagnosing accurately, stratified by clinical scenario, was performed. The combined improvement in diagnostic accuracy with early support over control was 13% (95% CI 7%-18%) but there was significant heterogeneity (I^2 53%, $p=0.004$). The combined improvement with late support over control was 7% (95% CI 2%-13%) but again there was significant heterogeneity (I^2 40%, $p=0.04$). The results are therefore reported separately for the UK and Greek samples.

Table 2 shows the mean diagnostic accuracy per experimental condition separately for the UK and Greek GP samples.

Experimental Condition	UK sample (N=297)	Greek sample (N=150)
Control	0.63 (0.19)	0.60 (0.19)
Early Support	0.69 (0.19)	0.71 (0.18)
Late Support	0.65 (0.18)	0.69 (0.16)

Table 2: Mean (SD) of correct diagnoses out of 9 per experimental condition for the UK and Greek samples.

Using independent samples t-test, a significant difference in diagnostic accuracy was detected in the UK sample between control and early support ($t=2.21$, $df=196$, $p=0.03$, two-tailed) but not between control and late support ($p=0.41$). In the Greek sample, both support conditions differed significantly from the control group: $t=3.13$,

df=98, p=0.002 for early support; t=2.65, df=99, p=0.009 for late support. There were no differences between the two support conditions in either sample.

The 6% absolute improvement with early support in the UK sample amounts to a number needed to treat (NNT) of 17: given cases of similar difficulty to those used in the study, one patient in 17, who might otherwise have been subject to misdiagnosis, would be correctly diagnosed if early support was used. In the Greek sample, the NNT with early support was 9 and with late support 11.

The differences in diagnostic accuracy between conditions were confirmed by multilevel logistic regression with random intercept. In the UK sample, the odds of diagnosing correctly were 1.31 times higher with early support than control (OR 1.31 [95% CI 1.04-1.66] p=0.023) and almost the same with late support as with control (OR 1.10 [95% CI 0.87-1.38] p=0.43). In the Greek sample, the odds of diagnosing correctly were 1.68 times higher with early support and 1.50 times higher with late support than control (OR 1.68 [95% CI 1.23-2.30] p=0.001 and OR 1.50 [95% CI 1.10-2.04] p=0.011). In the Greek sample, when the model accounted for scenario, the odds of diagnosing correctly with early support over control almost doubled (OR 1.93 [95% CI 1.30-2.89] p=0.001).

There were no differences between conditions in the amount of information gathered in either GP population (Table 3).

	Cues requested	
Experimental Condition	UK GPs (N=297)	Greek GPs (N=150)
Control	19.05 (5.11)	20.94 (5.83)
Early Support	19.62 (4.62)	20.67 (6.02)

Late Support	19.72 (4.95)	21.45 (5.91)
---------------------	--------------	--------------

Table 3: Means (SD) of information gathered by experimental condition for UK and Greek GPs.

4. Discussion

This randomised controlled study of early vs. late diagnostic support obtained a statistically significant improvement in the diagnostic accuracy of GPs in two different European countries (UK and Greece) by reminding them of possible diagnoses to consider early on in the clinical encounter. In Greece, late support also improved diagnostic accuracy over control significantly. The 6% absolute improvement in the UK sample may seem modest, however it amounts to a number needed to treat (NNT) of 17 for cases of similar difficulty, which is clinically important. Furthermore, it is the same as the net improvement in diagnostic accuracy found in a study that evaluated two existing diagnostic support systems, ILIAD and QMR, in a before-after design with hospital internists.¹⁹

Given that early support triggered by the reason for encounter was effective in both countries, our recommendation would be to prioritise the development of this type of DSS prototype. Not only early support would be simpler and technically easier for EHR vendors than late support (it would not require a system that monitor a physician's information gathering); it would also require a smaller change in the way GPs interact with their record: recording a reason for encounter at the start of the consultation. Late support, on the other hand, would require GPs to record on their EHR all the information gathered during the consultation and this may prove a significant hindrance for DSS adoption. However, early support will have its challenges too: recording a reason for encounter at the start of the consultation may sometimes prove difficult or impossible due to the unstructured and variable way patients present information. Patients may not offer the main reason for encounter at the start or may offer several reasons for encounter. The system needs to be flexible enough to cope with such variability and provide appropriate advice depending on any amount of information recorded at any point during the consultation. Therefore,

the system should be capable of providing advice with one or multiple pieces of information recorded at the start of the consultation or at any point thereafter.

We did not test a specific DSS, which allows our conclusions to be generalisable. Nevertheless, we followed certain guiding principles for designing the support in the study: that the support is automated, i.e., does not rely on physician request; that it is triggered during the consultation; and that it is minimally invasive, i.e., it does not trigger during information gathering. Furthermore, a small number of specific design decisions were made in order to operationalise diagnostic support in the study. For example, the early list of diagnostic suggestions remained on screen for at least 20 seconds and participants had to confirm that they read it before proceeding. This was done to ensure that the list was indeed read. Finally, we chose not to present diagnoses in order of prevalence but randomise the order for each participant, thus avoiding situations where they may ignore diagnoses low on the list. Further work (WT 2.3 and 2.4) will determine the optimal features of the human-computer interface for an early support DSS and how best to integrate it with the EHR and any other decision support in use, e.g., prescribing.

5. Conclusions

Reminding GPs of diagnostic possibilities to consider early on in the consultation, before they start narrowing down on one or a small number of hypotheses, can significantly improve diagnostic accuracy across a range of diagnostic scenarios, without significantly increasing the amount of information gathered. In a sample of GPs in Greece, both early and late reminders improved diagnostic accuracy over control. We thus propose to develop a DSS prototype, integrated with the EHR, that will provide early reminders in response to the reason for encounter but that will be flexible enough to make diagnostic suggestions at other points during the consultation and with varying amounts of information

6. Next steps

WT 2.3 has now commenced, involving the detailed requirements elicitation of a system to implement early support. This requires integration with a single EHR system (InPS, <http://www.inps4.co.uk>) and the clinical workflow. WT 5.2b, due to commence shortly, will develop the software required to link the DSS with the diagnostic ontology (WT 4.3 - Ontology service for diagnostic evidence & WT 4.4 - Web-based evidence service). A summary of this work is provided in Deliverable 10.3 (3rd year periodic report).

References

1. Silk N. What went wrong in 1000 negligence claims. *Health Care Risk Report*, November 2000:13-16.
2. Gandhi TK, Kachalia A, Thomas EJ, Puopolo AL, Yoon C, Brennan TA, et al. Missed and Delayed Diagnoses in the Ambulatory Setting: A Study of Closed Malpractice Claims. *Ann Intern Med* 2006;145(7):488-96.
3. Graber ML, Kissam S, Payne VL, Meyer AND, Sorensen A, Lenfestey N, et al. Cognitive interventions to reduce diagnostic error: a narrative review. *Bmj Qual Saf* 2012;21(7):535-57.
4. Garg AX, Adhikari NKJ, McDonald H, Rosas-Arellano MP, Devereaux PJ, Beyene J, et al. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes - A systematic review. *JAMA* 2005;293(10):1223-38.
5. Berner ES. Clinical Decision Support Systems: State of the Art. AHRQ Publication No. 09-0069-EF. Rockville, Maryland: Agency for Healthcare Research and Quality, 2009.
6. Ramnarayan P, Cronje N, Brown R, Negus R, Coode B, Moss P, et al. Validation of a diagnostic reminder system in emergency medicine: a multi-centre study. *Emerg Med J* 2007;24(9):619-24.
7. Feldman MJ, Barnett GO. An approach to evaluating the accuracy of DXplain. *Comput Methods Programs Biomed* 1991;35(4):261-6.
8. Ramnarayan P, Kapoor R, Coren M, Nanduri V, Tomlinson A, Taylor P, et al. Measuring the Impact of Diagnostic Decision Support on the Quality of Clinical Decision Making: Development of a Reliable and Valid Composite Score. *J Am Med Inform Assoc* 2003;10(6):563.
9. Ramnarayan P, Roberts GC, Coren M, Nanduri V, Tomlinson A, Taylor P, et al. Assessment of the potential impact of a reminder system on the reduction of diagnostic errors: a quasi-experimental study. *BMC Medical Informatics and Decision Making* 2006;6 (22):Open Access.
10. Ramnarayan P, Winrow A, Coren M, Nanduri V, Buchdahl R, Jacobs B, et al. Diagnostic omission errors in acute paediatric practice: impact of a reminder system on decision-making. *BMC Medical Informatics and Decision Making* 2006;6:37.
11. Friedman CP, Gatti GG, Franz TM, Murphy GC, Wolf FM, Heckerling PS, et al. Do Physicians Know When Their Diagnoses Are Correct? Implications for Decision Support and Error Reduction. *J Gen Intern Med* 2005;20(4):334-39.
12. Mamede S, Schmidt H, Rikers R, Penaforte JC, Coelho-Filho JM. Influence of perceived difficulty of cases on physicians' diagnostic reasoning. *Acad Med* 2008;83(12):1210-16.
13. Kawamoto K, Houlihan CA, Balas A, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success *BMJ* 2005;330:765-68.
14. Roshanov PS, Fernandes N, Wilczynski JM, Hemens BJ, You JJ, Handler SM, et al. Features of effective computerised clinical decision support systems: meta-regression of 162 randomised trials. *BMJ* 2013;346:f657.

15. Elstein AS, Shulman LS, Sprafka SA. *Medical problem solving: An analysis of clinical reasoning*. Cambridge, MA: Harvard University Press, 1978.
16. Kostopoulou O, Russo JE, Keenan G, Delaney BC, Douiri A. Information distortion in physicians' diagnostic judgments. *Medical decision making : an international journal of the Society for Medical Decision Making* 2012;32(6):831-9.
17. Kostopoulou O, Devereaux-Walsh C, Delaney BC. Missing celiac disease in family medicine: the importance of hypothesis generation. *Med Decis Making* 2009; 29(3):282-90.
18. Kostopoulou O, Mousoulis C, Delaney BC. Information search and information distortion in the diagnosis of an ambiguous presentation. *Judgment and Decision Making* 2009;4(5):408-18.
19. Friedman CP, Elstein AS, Wolf FM, Murphy GC, Franz TM, Heckerling PS, et al. Enhancement of Clinicians' Diagnostic Reasoning by Computer-Based Consultation: A Multisite Study of 2 Systems. *JAMA: The Journal of the American Medical Association* 1999;282(19):1851-56.
20. Luchins AS. Mechanization in Problem Solving: The Effect of Einstellung. *Psychological Monographs* 1942;54(Whole No. 248).
21. Kostopoulou O, Oudhoff J, Nath R, Delaney BC, Munro CW, Harries C, et al. Predictors of diagnostic accuracy and safe management in difficult diagnostic problems in Family Medicine. *Med Decis Making* 2008;28(5):668-80.

Abbreviations

(in alphabetical order)

BMI	Body Mass Index
COPD	Chronic Obstructive Pulmonary Disease
DSS	Diagnostic Support Systems
EHR	Electronic Health Record
GP	General Practitioner
InPS	In Practice Systems
KCL	King's College London
NIHR	National Institute for Health Research
NNT	Number Needed to Treat
RCSI	Royal College of Surgeons of Ireland
UoC	University of Crete

Appendix

Example of a scenario (angina vs. musculoskeletal pain) with the list of pre-determined cues and the diagnostic suggestions.

NAME	Andrew Marchant
Reason for encounter	"I have a pain in my chest, doctor, like a heavy weight. I first felt it when I lifted a lot of weight at work, a week ago. But I still feel it, on and off, especially when I get tired."
Background information	60 years old Caucasian Height: 1.70 m Weight: 88 kg (BMI 30.44) Smoking status: Non-smoker Last BP: 140/95, taken a year ago. Past medical history: Early knee osteoarthritis Current medications: Paracetamol, topical ibuprofen Last consultation: Low back pain a year ago. He is an infrequent attender. Appearance: As he enters the room, you notice that he is overweight.
HISTORY CUES	
Occupation	I'm a removals man. I've got my own lorry and everything.
Pain - location	I feel the pain right behind the sternum.
Pain - exacerbating factors	I first felt the pain when I lifted a box from the back of the lorry and carried it for about 10 metres. Since then, I feel it every time that I lift some heavy weight or move a piece of furniture.
Pain - relieving factors	When I take a rest, it goes away slowly.
Pain - duration of episode	It lasts about 15 minutes each time it comes.
Pain - When did you last feel it?	It last happened yesterday afternoon when I was moving some furniture for work.
Orthopnoea (shortness of breath when lying down)	I don't have these problems.
PND (Paroxysmal Nocturnal Dyspnoea)	I sleep well. It doesn't bother me at night.
Cough	No, I don't have a cough.
Previous incidents	I don't remember having chest pain in the past.
Pain - Radiation	I feel the pain in the centre of my chest - it doesn't go anywhere else.
Dizziness	I've never felt dizzy.
Indigestion	Haven't noticed a connection with food. I don't tend to get indigestion either.
Loss of consciousness/fainting	I've never fainted.
Breathless with pain?	I do get breathless when I'm lifting heavy things.
Sweaty with pain?	Not sure really - I am always sweaty at work!
Pain with specific movement (chest/arms)	No, moving my arms or twisting doesn't bring on the pain.

Pain when you press?	No, it doesn't hurt when I press.
Pain with general movement (walking, running, etc.)	I drive from place to place, so I'm not generally doing much walking. Definitely don't run.
Pain when you go up stairs?	I'm not sure, I get the lift.
Pain with breathing	No, I don't get any pain when I breathe in.
Fatigue	I find that I get more tired recently and short of breath when I carry things in the houses. I've never had such problems in the past. Initially, I didn't worry, I put it down to too much work and carrying a lot of weight.
Normal symptomatic level	I'm usually alright, especially when I'm not tired.
Family History	Both my parents are alive. My father is 83 and quite fit but my mother suffers from high blood pressure and diabetes. She's 79.
Alcohol	I usually have a couple of beers or some whisky with friends, almost every evening.
Are you worried?	I'm worried enough to come and see you, doctor. This pain is persistent.
Social history	I'm married and I have 2 children.
Stress currently	I've had some financial problems. Work is slow with the economic crisis, as you know.
Pain - NOW	I don't feel it right now.
Self treatment	I haven't taken anything for the pain, I'm not a great one for tablets.
Palpitations	I haven't had any palpitations.
Ankle swelling	I haven't had any leg swelling.
EXAMINATION CUES	
Temperature	His temperature is 36.6 C.
Pulse	His pulse is 96 bpm and regular.
BP (Blood pressure)	His BP is 150/100.
Respiratory rate	His respiratory rate is 14 breaths per minute.
Wheeze in chest	There is no wheeze in the chest
Listen to chest	No abnormal findings
JVP (Jugular Venous Pressure)	JVP is not raised.
Listen to heart	No abnormal findings
Peripheral oedema	There is no peripheral oedema.
Chest examination	No chest wall tenderness on palpation. No pain on rotation of the thoracic spine.
Abdominal examination	No abnormal findings
INVESTIGATION CUES	
Resting ECG	The resting ECG is normal with sinus rhythm.
Cholesterol	Cholesterol testing not done.
GENERIC ANSWERS	
	No.
	No, I don't.
	No, I haven't.
	I have not had that.

	I have not had any problems with that.
	No abnormality.
	The test is normal.

List of diagnostic suggestions

- Angina
- Anxiety and mood disorders
- Aortic aneurysm
- Aortic valve regurgitation
- Aortic valve stenosis
- Biliary disease
- Diabetes
- GORD / Peptic ulceration
- Herpes zoster
- Hyperthyroidism
- LRTI, bacterial
- Musculoskeletal chest pain
- Pancreatitis
- Pericarditis
- Pulmonary embolism
- Rib fracture
- Viral illness