



Validation of a diagnostic reminder system in emergency medicine: a multi-centre study

Padmanabhan Ramnarayan, Natalie Cronje, Ruth Brown, Rupert Negus, Bill Coode, Philip Moss, Taj Hassan, Wayne Hamer and Joseph Britto

Emerg. Med. J. 2007;24:619-624
doi:10.1136/emj.2006.044107

Updated information and services can be found at:
<http://emj.bmj.com/cgi/content/full/24/9/619>

These include:

References

This article cites 20 articles, 11 of which can be accessed free at:
<http://emj.bmj.com/cgi/content/full/24/9/619#BIBL>

Rapid responses

You can respond to this article at:
<http://emj.bmj.com/cgi/eletter-submit/24/9/619>

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article

Notes

To order reprints of this article go to:
<http://journals.bmj.com/cgi/reprintform>

To subscribe to *Emergency Medicine Journal* go to:
<http://journals.bmj.com/subscriptions/>

ORIGINAL ARTICLE

Validation of a diagnostic reminder system in emergency medicine: a multi-centre study

Padmanabhan Ramnarayan, Natalie Cronje, Ruth Brown, Rupert Negus, Bill Coode, Philip Moss, Taj Hassan, Wayne Hamer, Joseph Britto



Information on the Isabel system can be found at www.isabelhealthcare.com

See end of article for authors' affiliations

Correspondence to: Padmanabhan Ramnarayan, 44 B Bedford Row, Children's Acute Transport Service, London WC1R 4LL, UK; ram@isabelhealthcare.com

Accepted 22 March 2007

Emerg Med J 2007;24:619–624. doi: 10.1136/emj.2006.044107

Background: Diagnostic error is a significant problem in emergency medicine, where initial clinical assessment and decision making is often based on incomplete clinical information. Traditional computerised diagnostic systems have been of limited use in the acute setting, mainly due to the need for lengthy system consultation. We evaluated a novel web-based reminder system, which provides rapid diagnostic advice to users based on free text search terms.

Methods: Clinical data collected from patients presenting to three emergency departments with acute medical problems were entered into the diagnostic system. The displayed results were assessed against the final discharge diagnoses for patients who were admitted to hospital (diagnostic accuracy) and against a set of "appropriate" diagnoses for each case provided by an expert panel (potential utility).

Results: Data were collected from 594 patients (53.4% of screened attendances). Mean age was 49.4 years (95% CI 47.7 to 51.1) and the majority had significant past illnesses. Most were assessed first by junior doctors (70%) and 266/594 (44.6%) were admitted to hospital. Overall, the diagnostic system displayed the final discharge diagnosis in 95% of inpatients and 90% of "must-not-miss" diagnoses suggested by the expert panel. The discharge diagnosis appeared within the first 10 suggestions in 78% of cases.

Conclusions: The Isabel diagnostic aid has been shown to be of potential use in reminding junior doctors of key diagnoses in the emergency department. The effects of its widespread use on decision making and diagnostic error can be clarified by evaluating its impact on routine clinical decision making.

Emergency departments (EDs) have been shown to be high risk clinical areas for the occurrence of medical adverse events.^{1,2} In contrast to the inpatient setting where diagnostic delays and missed diagnoses account for only 10–15% of adverse events, diagnostic errors are more frequent and assume greater significance in primary care, emergency medicine and critical care, resulting in a large number of successful negligence claims.^{3–6} Various reasons may account for this difference. In EDs, acute clinical presentations are often characterised by incomplete and poor quality information at initial assessment leading to considerable diagnostic uncertainty. Systematic factors such as frequent shift changes, overwork, time pressure and high patient throughput may further contribute to diagnostic errors.⁷ Cognitive biases inherent within diagnostic reasoning have also been shown to play a crucial role in perpetuating patient safety breaches during diagnostic assessment.^{8,9} Despite easier access to health information for the public through NHS Direct and NHS Online, the demand for emergency care is steadily rising,¹⁰ and it is possible that greater pressure on emergency staff to cut waiting times and increase efficiency with limited resources will lead to a higher incidence of patient safety incidents during diagnostic assessment.

The use of computerised diagnostic decision support systems (DDSS) has been proposed as a technological solution to minimise diagnostic error.¹¹ A number of DDSS have been developed over the past few decades, but most have not shown promise in EDs, either because they focussed on a narrow clinical problem (eg, chest pain) or because computerised aids intended for general use were used only for consultation in rare and complex diagnostic dilemmas. In order to enter complete clinical data into these systems using system-specific medical terminology, considerable user motivation and time was required (often 20–40 min of data entry time).¹² Isabel

(www.isabelhealthcare.com) is a novel DDSS which was primarily developed for acute paediatrics. Isabel users enter their search terms in natural language free text, and are shown a list of diagnostic suggestions (up to a maximum of 30, displayed on three consecutive pages, 10 diagnoses per page), which are arranged by body system (eg, cardiovascular, gastrointestinal) rather than by clinical probability.^{13,14} These suggestions are intended only as reminders to prompt clinicians to consider them in their diagnostic investigation, not as "correct" choices or "likely" diagnoses. The system uses statistical natural language processing techniques to search an underlying knowledge base containing textual descriptions of >10 000 diseases. In clinical trials performed in acute paediatrics, mean Isabel usage time was <3 min, and the system reminded junior doctors to consider clinically significant diagnoses in 12.5% of cases.¹⁵ The DDSS was extended to cover adult disease conditions in January 2005.¹⁶

Although the extended Isabel system was closely modelled on the paediatric version, a large scale validation of the newly developed system was felt necessary to establish its diagnostic accuracy and potential utility for a range of clinical scenarios among adult patients in EDs. This was especially important since acute presentations in adults differ significantly from those in children. Adults may have multiple pre-morbid conditions that confound their acute illness. Further, the spectrum of diseases encountered is quite distinct, and the relative importance of diagnoses during initial assessment may be different from that in paediatric cases. It was also postulated that the greater amount of clinical detail available on adults presenting to EDs may prolong data entry time. In addition, since the Isabel system relies on extracting key concepts from

Abbreviations: DDSS, diagnostic decision support systems; ED, emergency department; RA, research assistant

Table 1 Characteristics of participating emergency departments

	Centre A	Centre B	Centre C	Overall
Nature	Teaching	Teaching	Teaching	
Annual attendances (majors)	74 000	80 000	44 000	198 000
Inpatient admission rate	16%	29.8%	19%	21.6%
Clinical decision unit	Yes	Yes	No	
Number of consultants	5	5	4	

its knowledge base (textbooks and journal articles), variations in natural language textual patterns in the medical sources used for the adult system may significantly influence its diagnostic suggestions.

METHODS

This preliminary validation study was designed purely to examine the clinical performance of the Isabel system and identify its potential utility in adult EDs, not to assess its impact on clinical practice or diagnostic errors. Therefore, clinicians were not allowed real-time access to the system during patient assessment in this study. Clinical data from ED patients presenting with a range of acute medical problems were used to validate the results of the DDSS. The study was approved by the London multi-regional ethics committee (04/MREC02/41) and relevant local research governance committees.

Study centres

A convenience sample of four UK EDs were selected for data collection during the study. Due to significant delay in completing research governance procedures at one centre, data were finally only collected from three participating sites. The characteristics of the study sites are summarised in table 1.

Study patient data

Data were collected from all consecutive patients over 16 years old presenting to the "majors area" or resuscitation rooms with an acute medical problem. Patients presenting to "minors" or a similar area, patients with surgical complaints (including trauma, orthopaedics, ENT, ophthalmology and gynaecology), post-operative surgical problems, psychiatric problems (including substance and alcohol abuse) and complaints directly related to pregnancy were excluded. Patients presenting for reassessment of the same clinical problem within a week of an earlier visit to the ED were also excluded.

Study procedure

This study was a prospective, multi-centre observational study utilising medical case note review. No interventions were performed on patients.

Screening for eligibility

The attendance records of all consecutive patients presenting to an ED "majors area" within a pre-designated 2-week period were screened for eligibility by the primary research assistant (RA). Screening was performed at study sites one after the other, ie, the 2-week period was different for each centre. The complete medical and nursing notes of patients with presenting complaints that fitted study criteria were selected for further review. Reasons for exclusion were recorded for the remainder using specific codes established a priori. Where the RA was unsure of study eligibility, patients were included and data were collected. Such notes were reviewed at regular intervals by the principal investigator, and a final decision regarding study eligibility was made. To assess inter-rater reliability during

Table 2 Template of patient summary report provided to panel

Patient characteristics
Age, gender, ethnic origin
Clinical presentation
Registering complaint(s)
Triage category
Vital signs at triage
Presenting clinical symptoms
Relevant co-morbidities and family history
Current medications
Positive clinical signs
Initial tests performed and results (if available)

screening for eligibility, a second RA examined patient notes from two randomly chosen dates within the specified 2-week period at each study ED. The primary RA was blinded to these dates. Concordance between the two RAs was calculated using the kappa statistic (κ 0.65, 95% CI 0.62 to 0.68).

Data collection

From eligible patient notes, the primary RA extracted data regarding patient details; date and time of patient assessment; details of clinical presentation such as symptoms, past medical and family history; examination findings, tests performed and results available at the end of complete assessment by the first examining clinician; differential diagnosis and management plan of the first examining clinician; referral for specialist or senior opinion; and outcome of ED assessment. This was entered directly into an Access database (Microsoft, Reading, UK) by means of pre-designed electronic forms. During data entry, both positive and negative symptoms, signs and test results were collected as recorded in the patient notes. Following complete data collection at all centres, final discharge diagnoses for patients at the end of ED assessment, recorded either on discharge letters or on ED electronic systems, were ascertained. For patients admitted as inpatients, final primary diagnoses at hospital discharge were obtained from hospital electronic coding systems.

Data quality assurance was achieved by multiple means including: use of a training manual created before study commencement containing standardised case examples to practise screening, data collection and abstraction; the use of 25 medical records to practise data collection, with doubts being clarified by the study investigator at study outset; weekly meetings to discuss data abstraction issues and examine collected data; and a log of discussions for ready reference. Reliability of the data collection process was established by randomly assigning 50% of eligible patient notes from two randomly chosen dates to a second RA. Concordance for key variables was analysed using the kappa statistic (κ 0.58, 95% CI 0.52 to 0.64).

Expert panel

An expert panel was set up at each study site consisting of two consultants (attending physicians), which met regularly to provide gold standard diagnoses for a randomly selected subset of study patients. At each panel meeting, moderated by the primary RA, data collected during initial ED assessment for each patient were provided in the form of a pre-formatted clinical summary report generated from the Access database (table 2).

Presenting clinical symptoms were provided as recorded in the patient notes. Only relevant co-morbidities, family history, positive clinical signs and results of initial tests (if performed by the examining clinician) were provided. The panel were blinded

Figure 1 Clinical data extracted from patient medical notes is entered into the diagnostic reminder system in free text natural language. Filters for age, gender, pregnancy status and geographical region are selected as drop-down choices.

to the ED where the patient was assessed, clinical decisions made and eventual patient outcome. The panel were instructed to provide, for each case following mutual consensus, a set of “must-not-miss” diagnoses, defined as key diagnoses that would influence patient management, ie, result in a specific action, either eliciting further history or physical examination or initiating new tests and/or treatments. To establish concordance between the three expert panels, a random selection of 5% of study patients was assigned to each panel in blinded fashion. For this subset, gold standard diagnoses were defined as those suggested by two or more panels.

DDSS data entry

Concurrent with the data collection process, an Isabel prototype was created from the original paediatric version. In order to generate a mature DDSS for use in adult patients, this prototype needed refinement. A total of 130 notes randomly drawn from patients not admitted to hospital were used for this final

fine-tuning (development set). In an iterative process, the system’s results for each case were critically examined by clinicians in the development team. To generate a focussed list of diagnoses, each diagnosis in the Isabel database had to be tagged to particular age group(s), gender and regions in which the disease was commonly seen (paediatric tags could not be used for adult patients). Once a fully developed Isabel system was available, the remainder of the cases (464/594) were used to test its performance (validation set).

During the validation stage, the RA exported the clinical summary report (as presented to the panel) for each case from the Access database in the form of individual text files. Since the Isabel system accepted search terms only as text and negative findings could not be searched within its medical content, information from the patient summary report needed modification by the RA during data entry into Isabel. Patient characteristics (eg, age and gender) were input using a drop-down menu, numerical values from vital signs and test results were converted

Figure 2 Diagnostic reminders are grouped under body system. Ten suggestions are displayed on the first page with an option to view more diagnoses. Clicking RD leads directly to other Related Diagnoses.

Table 3 Breakdown of patients screened and reasons for exclusion

	Centre A	Centre B	Centre C	Total	p Value
Patient notes screened	344	403	366	1113	
Eligible patient notes (%)	200 (58.1)	194 (48.1)	200 (54.6)	594	0.02
Exclusions (%)	144	209	166	519	
Trauma	50 (34.7)	96 (45.9)	47 (28.3)	193	0.5
Surgical/post-surgical	15 (10.4)	10 (4.8)	20 (12.0)	45	0.03
Pregnancy-related	18 (12.5)	14 (6.7)	25 (15.0)	57	0.03
Psychiatric/psychosomatic	13 (9.0)	12 (5.7)	12 (7.2)	37	0.5
Substance abuse	15 (10.4)	23 (11.0)	21 (12.6)	59	0.8
Orthopaedic	2 (1.4)	19 (9.1)	5 (3.0)	26	0.001
Incomplete data	31 (21.5)	35 (16.7)	36 (21.7)	102	0.4

to text terms (eg, temperature 38.8°C into “fever”), and only positive findings and test results (when available) were entered. This procedure was standardised prior to data entry by establishing normal ranges for vital signs and test results. The patient summary was aggregated into a single block of text with each symptom, positive clinical finding, salient past illness and the result of each test placed on a separate line, and then pasted into the search box. Any amount of clinical information could be entered, although it was accepted that specificity of search results would improve with detailed data entry. Figure 1 illustrates this procedure. In fig 2, results as displayed in Isabel are shown.

Outcome measures

Two separate outcome measures were assessed. Diagnostic accuracy was used to provide an indication of the system's clinical performance and was defined as the proportion of cases admitted to hospital in which the final discharge diagnosis appeared among the DDSS suggestions. However, this did not provide an indication of the system's utility in terms of guiding appropriate patient management in clinical practice. Therefore, the proportion of cases in which the DDSS included the entire set of key diagnoses that were deemed as “must-not-miss” by the expert panel indicated its utility in an ED setting.

Table 4 Patient characteristics

	Centre A	Centre B	Centre C	Overall	p Value
Average age (years)	47.57	50.5	50.23	49.4	0.32
Male/female ratio	1.06	0.83	1.08	0.99	
Mode of arrival (%)					
Self	95 (47.5)	98 (50.5)	101 (49.9)	294 (49.5)	
Ambulance	100 (50.0)	96 (49.5)	96 (48.0)	292 (49.2)	0.92
Other	5	0	3	8	
Triage category (%)					
Unknown	7	0	2	9	
1	4	3	4	11	0.93
2	30 (15.1)	18 (11.4)	53 (26.9)	101	<0.001
3	110 (55.6)	98 (62.0)	97 (49.2)	305 (55.0)	0.41
4	44 (22.2)	38 (24.0)	41 (20.8)	123	0.83
5	3	1	0	4	
Examining clinician (%)					
SHO (resident)	162 (81.0)	154 (79.4)	129 (64.5)	445 (74.9)	0.001
Registrar (fellow)	23	19	22	64	
Consultant (attending)	2	4	2	8	
Unclear/other	13	17	47	77	
Time of assessment (%)					
0800–1800	118 (59.0)	126 (64.9)	111 (55.5)	355 (59.8)	
1800–0800	82	68	89	239	0.15
Significant co-morbidities	180	180	183	543	0.61
Significant family history	29	10	32	71	0.001
Primary assessment outcome (%)					
Admit	19 (9.5)	80 (41.2)	29 (14.5)	128	<0.001
Review after tests	2	7	3	12	
Senior opinion	18	9	19	46	
Specialist opinion	63	23	63	149	
Discharge	81	72	82	235	
Unclear/other	17	3	4	24	
Final outcome (%)					
Hospital admission	80 (40.0)	101 (52.0)	85 (42.5)	266 (44.8)	0.16
Discharge (no follow-up)	45	90	59	194	
Discharge (GP follow-up)	43	NR	38	81	
Discharge (OP follow-up)	13	NR	9	22	
Transferred out	6	NR	2	8	
Other/unclear	13	3	7	23	
Total	200	194	200	594	

GP, general practitioner; NR, not recorded; OP, out patient.

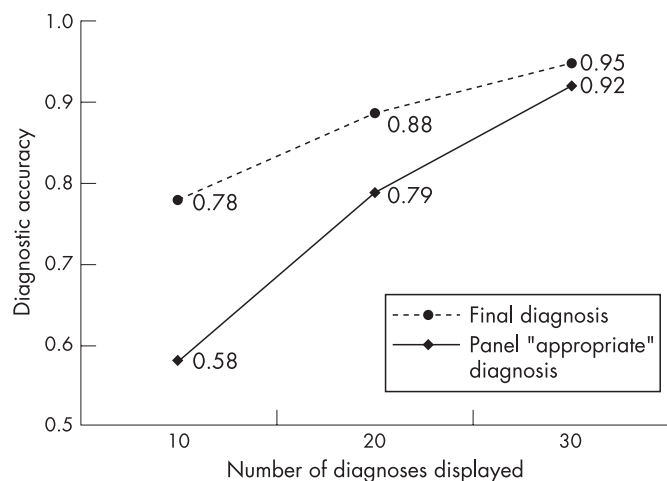


Figure 3 Diagnostic accuracy plotted for final diagnosis as well as expert panel "must-not-miss" diagnoses using 10, 20 and 30 results.

Sample size and statistical analysis

Using an estimated diagnostic accuracy of 85% (from the paediatric system) and an acceptable error rate of 3.5%, data were required from 450 patients to ensure adequate power. Differences between study EDs were analysed using the χ^2 test for proportions and ANOVA for continuous variables. Statistical significance was set at p value <0.05.

RESULTS

During the study period, 1113 consecutive patient notes were screened for study eligibility at the three centres. A total of 489 patients were excluded by the RA and a further 30 patients were judged as ineligible on secondary review (overall 46.6%). Therefore, 594 medical notes were reviewed in detail. There were significant differences between centres with respect to the proportion of patient notes excluded and reasons for exclusion (table 3).

Patient characteristics are summarised in table 4.

The mean age of patients included in the study was 49.4 years (95% CI 47.7 to 51.1) with an equal male to female ratio. A significant number of eligible patients were brought into EDs by ambulance (49.2%), and the majority were triaged into level 3 (55%). Most patients were seen by an SHO in the first instance (74.9%), and 40% were seen out of working hours (1800–0800). The majority of patients had past medical illnesses of note (91.4%). In addition, significant family history was present in a number of patients (12%). The primary examining clinician indicated having sought senior opinion in 7.7% and specialist opinion in nearly 25% of patients. Overall, 44.8% were admitted to hospital as inpatients, of whom 33% were discharged without further follow-up arrangements.

Diagnostic accuracy was measured using 217 inpatient discharge diagnoses available from 266 admissions. Overall, the DDSS displayed 206/217 diagnoses, with an accuracy rate of 95% (CI 92% to 98%). Seventy eight per cent of the discharge diagnoses were displayed on the first page (first ten suggestions).

Panel members examined 129 cases to provide "must-not-miss" diagnoses. A total of 30 cases were assessed by all three expert panels and 99 others by a single panel. In the former set, 52 diagnoses were suggested (mean 1.7 per case). Isabel displayed 50/52 suggestions (96%) in the list of its diagnostic reminders, with the majority present on the first page (35/50, 70%). For the latter set, 100 "must-not-miss" diagnoses were provided by the panel (mean 1 per case). Isabel displayed 90/100 suggestions; 53/100 were present on the first page (first 10

Table 5 Case assessment and "must-not-miss" diagnoses

Patient characteristics
69 years old, male, white, Irish
Clinical presentation
Registering complaint(s): abdominal pain
Triage category: 3
Vital signs at triage: temp 38.8, BP 108/65 mm Hg, heart rate 120/min, respiratory rate 20 bpm, saturations 96% in room air, GCS 15/15
Presenting clinical symptoms
Fever
Rigors
Right iliac fossa pain radiating to central abdomen
Vomited
Nausea
Chronic blood per rectum
Reduced appetite
Constipated
Passing flatus
Chronic productive cough
Relevant co-morbidities and family history
Peripheral vascular disease
Diabetes mellitus
Ischaemic heart disease
Smoker
Current medications
Simvastatin
Metformin
Positive clinical signs
Bibasal crackles
Abdominal tenderness with percussion
Tender rectum
Pallor
Dehydrated
Fever
Tachycardia
Initial tests performed and results (if available)
Leucocytosis
High serum bilirubin
Elevated CRP
Sinus tachycardia on ECG
Thickened right colon, suspected mass on ultrasound scan abdomen
Panel assessment – "must not miss" diagnoses to consider
Acute appendicitis
Carcinoma colon
Ischaemic bowel

reminders). Comprehensiveness improved significantly when all three pages were examined (fig 3). An example of one expert panel's assessment of a case and relevant "must-not-miss" diagnoses is shown in table 5.

DISCUSSION

We have demonstrated in this large validation study that the Isabel diagnostic aid demonstrates significant accuracy for hospital discharge diagnosis among patients presenting to EDs with acute medical problems. The system also showed potential utility in the ED setting by including all key diagnoses among its diagnostic reminders.

The Isabel system has been previously evaluated in acute paediatrics, and shown to display the final diagnosis on the first page in >90% cases drawn from real life practice.¹⁷ The heterogeneous nature of acute presentations among adult patients and their lengthy past history may have resulted in more complex clinical data entry accounting for some decrease in accuracy in this study. Most adult patients had significant past medical illnesses and were on numerous medications. In addition, textual descriptions of diseases in the current Isabel knowledge base may be quite different between adult and paediatric sources resulting in a poorer match between the clinical data entered and the diagnostic results generated. Despite these findings, Isabel's diagnostic performance is better than that of previously described

DDSS.¹⁸ VanScoy *et al* showed in an ED setting using two diagnostic systems, QMR and ILIAD, that the final diagnosis was present in the top five choices in only 30% of cases. Data entry time was prolonged since detailed case descriptions needed to be matched to system-specific terminology and specific training was required in the use of the DDSS.¹⁹ In our study, clinical assessment was entered into Isabel in the examining clinicians' own words in free text, leading to the conclusion that rapid and easy use of the system is possible without prolonged data entry. In this context, Graber *et al* have also recently shown that merely pasting the entire history and physical examination section from the Case Records of the Massachusetts General Hospital series as text without any modification into Isabel resulted in a diagnostic accuracy rate of 74%.²⁰

Identifying the precise patient population in which DDSS might prove useful is a major challenge. Expert systems such as QMR were intended to be used in a diagnostic dilemma. However, there is sufficient evidence that diagnostic errors occur during routine practice, and that there is poor correlation between physicians' diagnostic accuracy and their own perception of the need for diagnostic assistance.²¹ We chose to focus on validating Isabel against a pre-selected subset of ED patients (acute medical problems seen in the "majors area"). Nearly half of all patients screened qualified using our liberal study criteria, although it is improbable that in practice medical staff would have used Isabel in all these patients. Experience from our paediatric clinical study indicates that clinicians may seek diagnostic advice in 5–7% of acute medical presentations (three to five ED patients per day).²² Identifying the optimal parameter against which DDSS performance can be measured also remains controversial.²³ We used hospital discharge diagnoses and an expert panel's opinion of key diagnoses to provide a combined view of system accuracy and utility. During this initial evaluation, we deliberately denied clinicians access to Isabel. Yet, by extrapolation from these results, it seems likely that clinicians will use and benefit from its diagnostic advice in situations of uncertainty, especially since minimal data entry time was required. Integration of the DDSS into an electronic medical record may allow active diagnostic advice to be delivered to staff with minimal effort. Such an interface has been developed recently.²⁴

LIMITATIONS

The main limitation of this study was the fact that users did not interact with the system making it difficult to estimate its true utility in practice. The impact of a DDSS is best measured by its ability to improve clinicians' diagnostic assessment; in addition, unexpected negative effects might be seen. We used electronic systems or discharge summaries to provide data on final diagnoses, but these sources have been shown to be unrepresentative and of variable quality. However, due to logistical reasons, we could not follow up all patients seen in this study. Also, as ED discharge diagnoses on patients were also missing in a number of patients, we used inpatient discharge diagnoses for our main outcome analysis.

CONCLUSIONS

Diagnostic assistance may be useful in a large proportion of patients seen in an emergency department. The Isabel diagnostic aid performs with an acceptable degree of clinical accuracy in this setting. Further studies to elucidate its effects on decision making and diagnostic error are essential in order to clarify its role in routine practice.

ACKNOWLEDGEMENTS

We gratefully acknowledge the advice during study design provided by Dr Paul Taylor.

Authors' affiliations

Padmanabhan Ramnarayan, Children's Acute Transport Service, London, UK

Natalie Cronje, Isabel Healthcare, London, UK

Ruth Brown, Rupert Negus, St Mary's Hospital, London, UK

Bill Coode, Philip Moss, Newham General Hospital, London, UK

Taj Hassan, Wayne Hamer, Leeds General Infirmary, Leeds, UK

Joseph Britto, Isabel Healthcare, Reston, VA, USA

Funding: This study was supported by a research grant from the National Health Service (NHS) Research & Development Unit, London. The sponsor did not influence the study design; the collection, analysis, and interpretation of data; the writing of the manuscript; or the decision to submit the manuscript for publication.

Competing interests: The Isabel system is currently managed by Isabel Healthcare, and is available only to subscribed individual and institutional users. Dr Ramnarayan is a part-time research advisor for Isabel Healthcare, Ms Cronje was employed as a research assistant by Isabel Healthcare for this study, and Dr Britto is Clinical Director of Isabel Healthcare. All other authors declare that they have no competing interests.

REFERENCES

- 1 Weingart SN, Wilson RM, Gibberd RW, *et al*. Epidemiology of medical error. *BMJ* 2000;**320**(7237):774–7.
- 2 Burroughs TE, Waterman AD, Gallagher TH, *et al*. Patient concerns about medical errors in emergency departments. *Acad Emerg Med* 2005;**12**(1):57–64.
- 3 Leape L, Brennan TA, Laird N, *et al*. The nature of adverse events in hospitalized patients. Results of the Harvard Medical Practice Study II. *N Engl J Med* 1991;**324**:377–84.
- 4 Davis P, Lay-Yee R, Briant R, *et al*. Adverse events in New Zealand public hospitals II: preventability and clinical context. *N Z Med J* 2003;**116**(1183):U624.
- 5 Sandars J, Esmail A. The frequency and nature of medical error in primary care: understanding the diversity across studies. *Fam Pract* 2003;**20**(3):231–6.
- 6 Bartlett EE. Physicians' cognitive errors and their liability consequences. *J Healthc Risk Manage* 1998;**18**:62–9.
- 7 Driscoll P, Thomas M, Touquet R, *et al*. Risk management in accident and emergency medicine. In: Vincent CA, ed. *Clinical risk management. Enhancing patient safety*. London: BMJ Publications, 2001.
- 8 Croskerry P. The importance of cognitive errors in diagnosis and strategies to minimize them. *Acad Med* 2003;**78**(8):775–80.
- 9 Kovacs G, Croskerry P. Clinical decision making: an emergency medicine perspective. *Acad Emerg Med* 1999;**6**(9):947–52.
- 10 Bache J. Emergency medicine: past, present, and future. *J R Soc Med* 2005;**98**(6):255–8.
- 11 Graber M, Gordon R, Franklin N. Reducing diagnostic errors in medicine: what's the goal? *Acad Med* 2002;**77**(10):981–92.
- 12 Lemaire JB, Schaefer JP, Martin LA, *et al*. Effectiveness of the Quick Medical Reference as a diagnostic tool. *CMAJ* 1999;**161**(6):725–8.
- 13 Ramnarayan P, Tomlinson A, Kulkarni G, *et al*. A novel diagnostic aid (ISABEL): development and preliminary evaluation of clinical performance. *Medinfo* 2004;**11**(Pt 2):1091–5.
- 14 Greenough A. Help from ISABEL for pediatric diagnoses. *Lancet* 2002;**360**(9341):1259.
- 15 Ramnarayan P, Roberts GC, Coren M, *et al*. Assessment of the potential impact of a reminder system on the reduction of diagnostic errors: a quasi-experimental study. *BMC Med Inform Decis Mak* 2006;**6**:22.
- 16 E-health insider. <http://www.e-health-insider.com/news/item.cfm?ID=1009> (accessed 9 July 2007).
- 17 Ramnarayan P, Tomlinson A, Rao A, *et al*. ISABEL: a web-based differential diagnostic aid for paediatrics: results from an initial performance evaluation. *Arch Dis Child* 2003;**88**(5):408–13.
- 18 Berner ES, Webster GD, Shugerman AA, *et al*. Performance of four computer-based diagnostic systems. *N Engl J Med* 1994;**330**(25):1792–6.
- 19 Graber MA, VanScoy D. How well does decision support software perform in the emergency department? *Emerg Med J* 2003;**20**(5):426–8.
- 20 Graber ML, Mathews A. Performance of a web-based clinical diagnosis support system using whole text data entry. <http://www.isabelhealthcare.com/info/independentq1.doc> (accessed 9 July 2007).
- 21 Friedman CP, Gatti GG, Franz TM, *et al*. Do physicians know when their diagnoses are correct? Implications for decision support and error reduction. *J Gen Intern Med* 2005;**20**(4):334–9.
- 22 Ramnarayan P, Winrow A, Coren M, *et al*. Diagnostic omission errors in acute paediatric practice: impact of a reminder system on decision-making. *BMC Med Inform Decis Mak* 2006;**6**(1):37.
- 23 Berner ES. Diagnostic decision support systems: how to determine the gold standard? *J Am Med Inform Assoc* 2003;**10**(6):608–10.
- 24 PatientKeeper. http://www.patientkeeper.com/success_partners.html (accessed 9 July 2007).