PRO*forma* guidelines and care pathways: performance in trials and future plans

John Fox and Vivek Patkar and Richard Thomson and Elizabeth Black.
Advanced Computation Laboratory, Cancer Research UK, London WC2A 3PX

Abstract. PRO*forma* is a language for modelling clinical processes, along with associated tools and methods for creating clinical decision support, care planning, workflow and other applications. Of the applications that have been built using the language some have been evaluated in primary healthcare settings (drug prescribing, referral of suspected cancer patients, genetic risk assessment) and others in specialist care of patients, including breast cancer, leukaemia and management of HIV+ patients. Eight of these trials have included quantitative evaluations on a variety of measures of quality and/or effectiveness of care, and all have shown significant positive effects. This paper provides an overview of these results and previews the CREDO project, which aims to extend the PRO*forma* method from focused decision-support and workflow applications to supporting large, multi-disciplinary care pathways, such as cancer care.

1. BACKGROUND

As demands for more and better patient care continue to grow the shortcomings of medical services are increasingly recognized. *To Err is Human*, the influential report from the US Institute of Medicine [1] focused attention on human factors in achieving consistently safe decision-making in difficult conditions. Quality and safety are now major issues with many governments and agencies pursuing policies to improve care at both individual and organizational levels (see www.openclinical.org). The difficulties for hard-pressed clinicians to deliver consistently high quality, safe care are also increasingly well documented. In cancer services, for example, the joint report by the UK Commission for Healthcare Improvement and Audit Commission on cancer care identifies a wide range of challenges to the consistent delivery of effective detection, diagnosis and treatment.

Decision support systems have an important contribution to make to this problem. In the systematic review of 100 trials of decision support systems by Garg et al [2] 62/97 studies demonstrated improved practitioner performance (64%). These broke down into 16/21 (76%) for simple reminder systems, 4/10 (40%) in diagnosis applications, 23/37 (62%) "disease management" and 19/29 (66%) drug dosing applications. Decision technologies are now seen as a key resource in the delivery of modern, evidence-based clinical services. For example the UK has committed £6.7 billion over 6 years as part of its modernization program for the National Health Service through it National Programme for IT in which the successful deployment of decision support is a key strategic objective [3].

The Garg et al review subsumes a range of techniques and approaches to improving decision-making. Despite the positive overall results of the study this is a problem for the field in that if decision support systems are to become generally accepted by clinical users we need to rationalize and ideally standardize the systems they are asked to use. Without standards it will also be difficult to achieve reusability, scalability, interoperability and other technical features that are needed to attract wide adoption [4]. The PRO*forma* approach to this problem proposes a standard, computer executable language and interchange format for modeling clinical decisions and processes [5]. PRO*forma* shares technical features with other proposals (see review by Peleg et al, 2002 [6]; a more recent overview of current approaches can be found at http://www.openclinical.org/gmmintro.html) but it is distinctive in being grounded in a general model of human decision-making and a formal framework for decision-making and planning with an associated formal semantics [7, 8].

Peleg et al [6] and de Clercq et al [9] have carried out independent comparisons of published approaches to modelling guidelines and protocols. PRO*forma* appears to be distinctive in having formal foundations for the language, notably the decision-making framework based on logical arguments rather than traditional decision analysis. The constructs of the language are principled and, experience suggests, intuitive for clinicians. The former characteristic supports the creation of reliable and sound applications, the latter helps to ensure that healthcare professionals are empowered to oversee application development and safe use in practice – neither of which are typical of conventional software.

The language has provided a foundation for a number of novel technologies and applications, and this has led to an encouraging body of evidence that the general approach has clinical value. This evidence is summarized below. The syntax and operational semantics of the language have been published in the open literature [10] to facilitate its use as an interchange format and to promote reusability of content and interoperability of applications and application components (see also [4]). Comprehensive application development tools are also available. Experience of their utility has now been accumulated with several such tools, including the Tallis authoring and web publishing suite which is available for research use (see web site at www.acl.icnet.uk/TallisTraining) and the commercial Arezzo® system available from *Infer*Med Ltd. (www.infermed.com).

2. REVIEW OF EVALUATIONS: DECISIONS AND GUIDELINES

The first practical application of PRO*forma* was CAPSULE, a system for assisting general practitioners in prescribing for common conditions [11]. Prescribing is an important application for decision support in many clinical areas, because of the amount of knowledge clinicians must absorb to stay up to date. CAPSULE took in patient data and suggested a short list of possible medications, together with the arguments for and against the different options. The trial showed potential

substantial improvements in decision accuracy (about 30%), improved resource use and faster decision-making.

LISA is an application for advising on dose adjustment in treatment of children with acute lymphoblastic leukaemia. A trial of LISA has shown that without decision support clinicians deviated from the trial protocol on 37% of occasions but with support this dropped to zero, and 35/36 of the clinicians said they would use LISA if it were routinely available [12]. Note that this trial is indicative rather than definitive since it was run on retrospective cases. LISA has now been incorporated into a prospective trial led by Cancer Research UK's Paediatric Oncology Unit at the London Hospital and integrated into *Infer*Med's MACRO® clinical trial manager.

An important area of application of decision support and care planning technologies is in clinical genetics. As more genetic markers become available, individuals who are themselves well but worried they are at risk for a particular disease because of family history seek guidance from their personal physicians and other non-specialists. Physicians are frequently willing to provide advice and support, but they do not generally feel they have the expertise to take or analyse a complex family history, and the problem of communicating risk to patients without statistical understanding is notoriously difficult. RAGs was developed to help the general practitioner take a family history, assess risk and explain risk factors to patients [13]. An application for breast and ovarian cancer risk was evaluated with Oxfordshire general practitioners, showing that with this support they produce more accurate family trees, risk assessments and management decisions [14]. In a comparison with other software and paper and pencil RAGs was chosen as the preferred tool 91.7% of the time.

If the initial risk assessment is confirmed by subsequent genetic testing then this opens the way to providing a high level of individualized care. However, customizing an individual's care is difficult and time-consuming, and it is likely to be difficult (and expensive) for busy clinical teams to implement personalized patient care plans. REACT is a treatment planning tool that provides the ability to personalize care plans taking into account individual risk information and patient's individual preferences [15]. A trial of the system to assist in care planning in a genetic counseling setting has produced encouraging results; 7/8 clinical geneticists were initially skeptical about its value but after completing the study in an actor-based scenario 7/8 were strongly supportive of its value [16]. Note that the trial version of REACT uses PRO*forma* concepts of plans and argument based decision-making but does not represent the care plan explicitly. Polyphony, an experimental version imports PRO*forma* models into the REACT planning environment where the planning and scheduling capabilities of REACT can be applied to a set of PRO*forma* care plans and their interactions.

The general field of clinical genetics has produced the most definitive trial of a PRO*forma* application to date, for tailoring drug selection and dose to an individual's genetic profile. The Retrogram® system was developed by *Infer*Med for Hoffman la Roche (www.retrogram.org). Retrogram advises on the use of anti-retroviral therapy for HIV+ patients and is believed to be in use with more than 250 clinicians worldwide. The HAVANA multicentre trial [17] has shown that availability of genotype information significantly improves clinicians' decision-making as measured by viral load, but with Retrogram providing genotype interpretation and decision support services the viral loads of many more patients' were reduced to the target level. As genetic profiling becomes increasingly important in patient care Retrogram could well be an important model for the future.

The CADMIUM imaging system was an early version of our language model in combining conventional image processing with automated interpretation of images and diagnosis [18]. CADMIUM was evaluated in a study in which radiographers with specialist training interpreted screening mammograms with and without decision support. The question was whether such systems could play a role in, or even substitute for consultant radiologists, in routine breast cancer screening. The system was designed to identify micro-calcifications in breast tissue automatically and interpret the pattern of calcifications in terms of whether they are likely to indicate benign or malignant abnormalities. The study indicated that the system significantly increased the rate of correct classifications of malignant and benign abnormalities while also reducing cancer misses and false positives.

3. SUPPORTING WORKFLOWS, COMPLEX CARE PATHWAYS

The applications described above, and the great majority of published studies of clinical decision-support systems, are concerned with supporting a single clinical decision. However, many routine clinical procedures involve multiple decisions, which are naturally carried out in a particular order dependent on the availability of data and test results, completion of procedures and conclusions about other decisions. Our current work is aimed at looking at whether PRO*forma* can model more complex workflows, where decisions are just some of the tasks required in the care plan. An example of this is TADS, a decision support system for the "Triple Assessment" of patients with suspected breast cancer. This typically takes place at a "same day" clinic where the aim is to achieve a provisional diagnosis and risk assessment of women with breast symptoms. Triple assessment includes:

1. Clinical Examination;
2. Various forms of imaging: (mammography, ultrasound, MRI,);
3. Histo-pathology: Fine needle aspiration or Core needle biopsy.

The assessments and decisions from these three activities feed in to a final management decision about whether she can be discharged or referred to the multidisciplinary team to oversee her treatment and follow-up. Figure 1 shows a high level representation of the tasks in this workflow, using the graphical representation of the Tallis authoring system. Here a diamond represents any kind of data acquisition (e.g. filling in a data entry form, retrieving information from an electronic patient record or capturing data from a medical device such as an imaging system). Decisions are represented as circles. There are 4 here: family history and risk assessment; a decision about what imaging studies to do if any, what biopsy method if any, and the final management decision. Figure 2 shows a typical web-browser display of a TADS decision, with the arguments for and against the different imaging options and links to relevant backing information.

Twenty four practising breast clinicians from UK National Health Service (NHS) hospitals participated in the study. A balanced-block crossover experiment was conducted using paper cases, followed by a questionnaire study and semi-structured interviews. At the decision points the clinicians made significantly more deviations from guideline recommendations without the support of the tool.

Ignoring minor deviations, 16 potentially critical errors arose in the no-decision-support arm of the trial compared with just one when decision-support was available. Opinions of participating clinicians towards the decision support tool became more positive after they had used it [19].

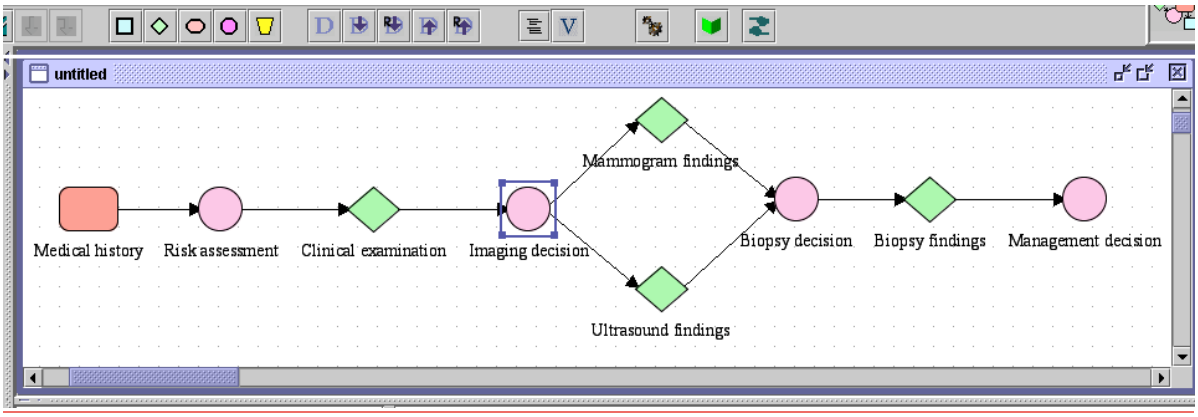

Figure 1: representation of triple assessment workflow. The decision nodes represented by circles are embedded at various points in the workflow

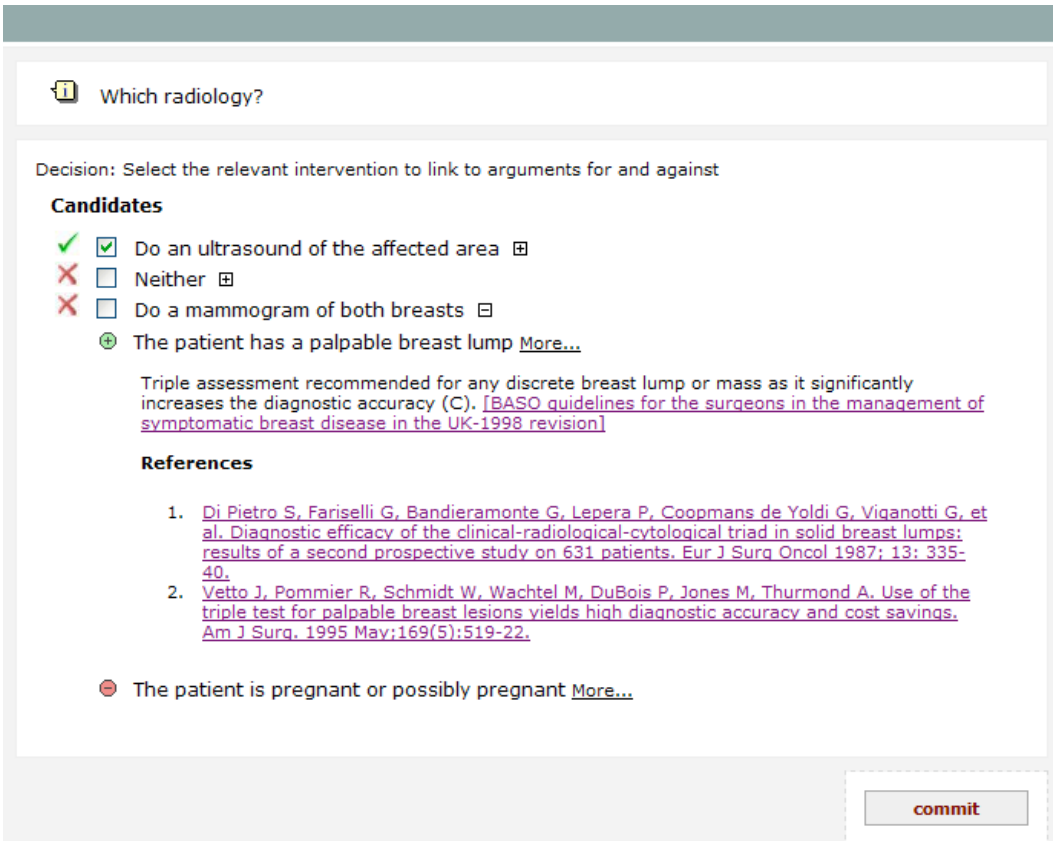

Figure 2: TADS support for the imaging decision after completing medical history and clinical examination tasks. The system recommends an ultrasound scan but recommends against mammography and doing nothing. For the decision option 'Do a mammogram of both breasts', arguments for and against have been expanded. Links are provided to the relevant supporting literature, which can be pulled down from the web (e.g. from PubMed) by the user if required.

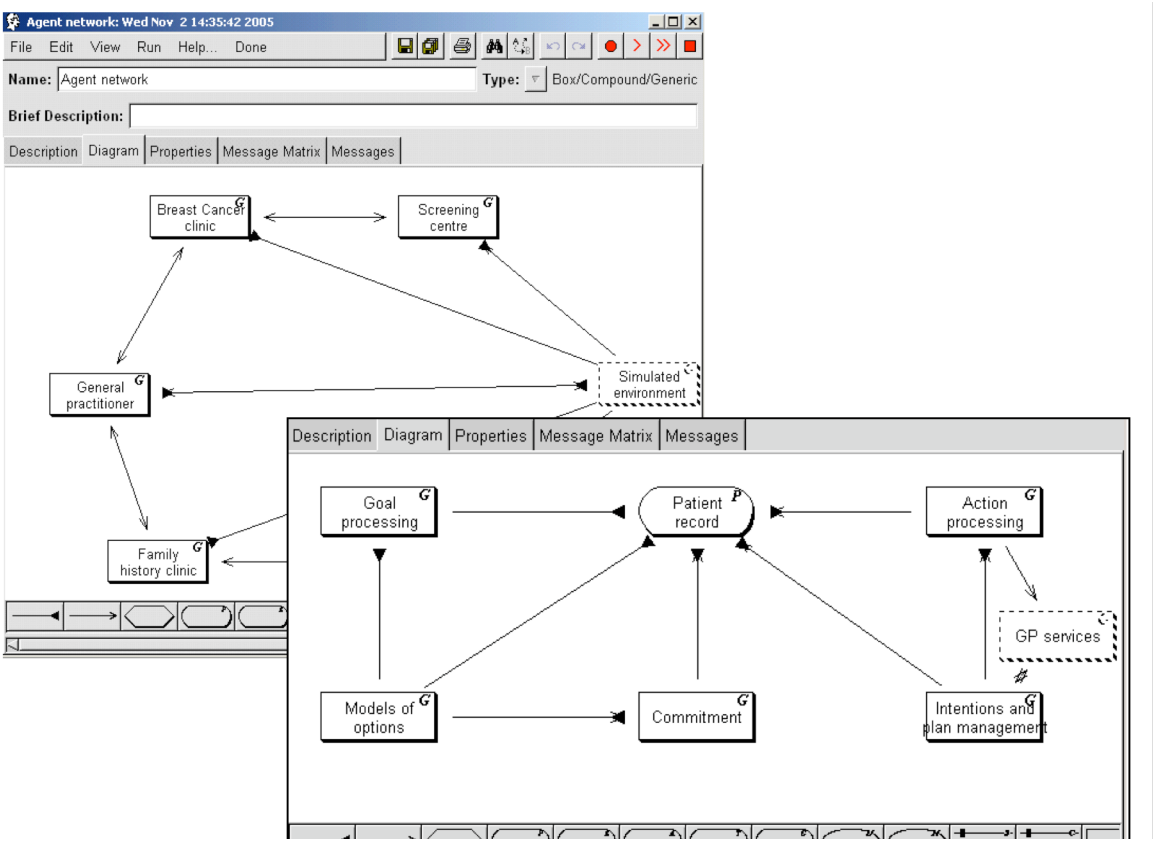

Figure 3: Black's (2002) agent network for simulating clinical referrals showing the architecture of one agent (inset). The component on the bottom right provides a store and execution function for PRO*forma* tasks.

To the best of our knowledge, this is the first evaluation of a decision support tool that supports multiple interdependent decisions in cancer care. The results of this initial trial suggest that the system can significantly enhance safety and consistency of patient care in breast cancer by promoting compliance with best practice.

Despite their relative complexity workflows like this represent only a tiny part of the services that are required for managing a complex condition like cancer, and in many cases patient care involves many different people in different places and with different skills. For this reason we have also begun to look at the question of whether PRO*forma* can support distributed applications, where each application can be viewed as an agent that can cooperate with other PRO*forma* agents. Huang et al [20, 21] developed an early demonstration of this approach in which each agent was capable of making decisions using an early version of the PRO*forma* decision model [22]. The technical adequacy of this approach was demonstrated in an application for distributed breast cancer. Each agent was capable of making decisions based on an argumentation approach [22, 23] but possessed different expertise and therefore needed to seek assistance in certain circumstances. For example a "general practitioner" agent has broad but shallow knowledge of the patient while a "cancer specialist" agent has deeper but narrower expertise.

More recently Black [24, 25] has extended this framework to investigate the potential value of agents to support distributed services by simulating an agent network on a realistic medical problem. Her model consists of three general practitioners (GP), two family history clinics and a genetics clinic (see figure 3, top left). Simulated patients would "present" to their GPs, who would decide whether the patient required referral for specialist advice or not. The simulated GP could make decisions and enact plans specified as PRO*foma* processes. If the GP's decision criteria indicated referral, then the agent arranges this by sending a message to the appropriate referral centre. Depending on the queuing strategy used in the simulation the agent may also book a *provisional appointment* for the patient at the genetics clinic. The family history agent works in a similar fashion, but it decides whether the patient needs referral to the genetics clinic. If not then any provisional appointment at the genetics clinic will need to be cancelled. Once patients have been seen in the genetics clinic their route through the system is complete.

The purpose of this study was to investigate how a number of different agents could interact to implement a complex, distributed decision model, and to determine whether different referral strategies could affect the performance of the agent network. Black considered 4 referral strategies. (1) Patients are simply seen in the order they arrive at each centre. (2) If the GP decides that the patient needs referring to family history, then at the same time they make a provisional appointment for the patient at the genetics clinic (two versions), and (3) as strategy 2 but appointments at the genetics units can be reassigned according to priority made by the family history agent based on a risk assessment decision.

Strategy	Av. time at GP	Av. time at family	Av. time at genetics	total cycles
No Strategy	33	101	256	380
Ad hoc	10	53	184	247
1st come, 1st served	26	85	173	284
Prioritised on risk	21	116	193	330

Table 1. Average patient queuing times (in cycles) for all patients under all strategies (simulated).

4. THE CREDO TRIAL

The CREDO project is intended to build on these results in order to seek definitive data on the value of advanced tools for decision support, care planning and distributed patient management in complex multidisciplinary care. The aims of the project are to trial a comprehensive suite of computer services to support all phases of care from first presentation and diagnosis through treatment and follow-up, and facilitate communication and coordination across the multidisciplinary team. The first task in this project has been to develop a comprehensive service model that would be amenable to proformalisation. This model is illustrated schematically in figure 4, but is described in detail at the project web site (www.acl.icnet.uk/credoweb) where a video that explains the project can also be found. The aim of the CREDO trial is to address the following questions:

- What is the baseline quality and safety of typical decision-making in the management of breast cancer?
- Can PRO*forma* decision support yield improved consistency, quality and safety of clinical decision making through the patient journey and what are the most promising decision points at which to intervene?
- Can such services be offered in a form that is acceptable to and valued by clinicians? Can they free clinicians from administrative and operational burdens, allowing them to concentrate on looking after patients?
- Can such support for clinicians and patients lead to improved experience for the cancer patient?

The trials described in the previous sections included a number of examples of decision-making in breast cancer, many of which form elements of the CREDO care pathway. Our service model for breast cancer care suggests that across the whole multidisciplinary pathway there are approximately 65 decision points where there is potential for errors that could significantly impact efficacy of treatment or safety of the patient (note that the key step of triple assessment includes just 4 of these decisions). It is interesting to note that with such a complex pathway if error rates in practice are about 1% on average then with 65 decisions taken independently only 50% of women would get perfect care on average. If error rates are 5% then only 3% of women will get perfectly safe care! Evidence from the above studies and other literature, notably the US report *To Err is Human* suggest that the actual frequency of deviations from evidence-based recommendations in routine decision-making could be considerably higher.

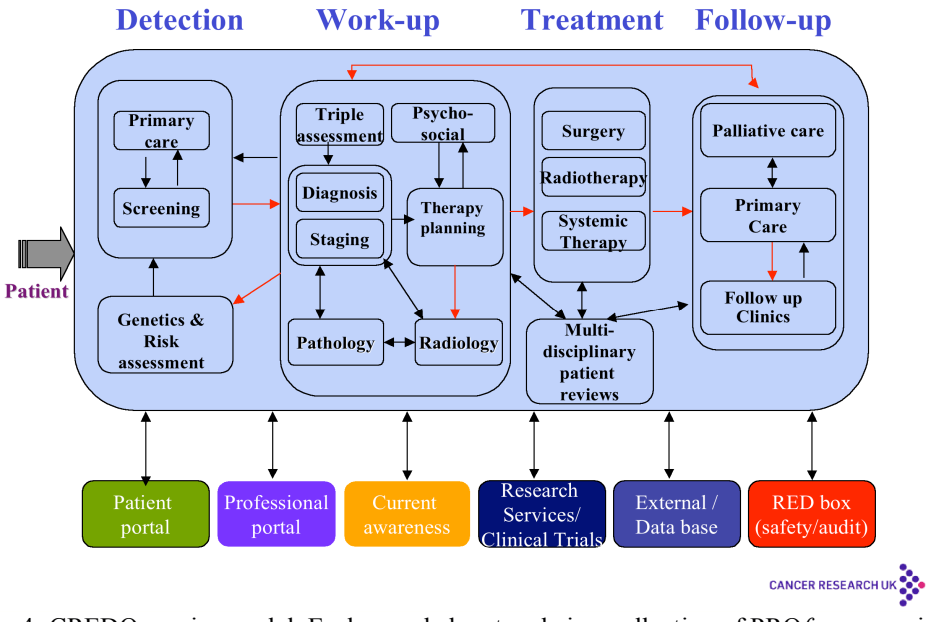

Figure 4: CREDO service model. Each rounded rectangle is a collection of PRO*forma* services. Black arrows = information flow, Red arrows = information flow + transfer of responsibility ("referral")

5. SUMMARY AND CONCLUSIONS

In eight studies of PRO*forma* applications where quantitative performance data have been obtained, all have shown significant positive results on a variety of techniques and measures. With the simplest assumption that the results were equally likely to go either way this pattern has a chance probability of less than 0.004 (0.50^8 = 0.0039) which is highly significant statistically. Current work is investigating the possibility that extended versions of PRO*forma* agents may be able to support distributed and multidisciplinary care with equally promising results.

6. REFERENCES

1. Kohn, et al, (editors) To Err Is Human: Building a Safer Health System, Committee on Quality of Health Care in America, Institute of Medicine, National Academy Press Washington, D.C. 1999
2. Garg A X, Adhikari N K J, McDonald H, Rosas-Arellano M P, Devereaux P J, Beyene J, Sam J, Haynes B, "Effects of Computerized Clinical Decision Support Systems on Practitioner Performance and Patient Outcomes: A Systematic Review" *J Am Med. Assoc*., Vol. 293 No. 10, March 9, 2005
3. Benson T and Page R "Care Pathway Modelling" Report of the National Programme for IT, 2005 (see http://www.openclinical.org/briefingpaperBenson.html)
4. Peleg M, Steele R, Thomson R, Patkar P, Rose T, Fox J "Open-Source Publishing of Medical Knowledge for Creation of Computer-Interpretable Guidelines" *Proc. 10th Conference on Artificial Intelligence in Medicine AIME05*, Berlin: Springer Verlag (2005)
5. Fox J, Johns N, Rahmanzadeh A, Thomson R "PRO*forma*: a method and language for specifying clinical guidelines and protocols", *Proc. Medical Informatics Europe*, Amsterdam: IOS Press 1996.
6. Das S, Fox J, Elsdon D, Hammond P, "A flexible architecture for a general intelligent agent" *Journal of Experimental and Theoretical Artificial Intelligence*, 9, 407-440, 1997.
7. Fox J and Das S, *Safe and Sound: Artificial Intelligence in Hazardous Applications*, jointly published by the American Association for Artificial Intelligence and MIT Press, Cambridge MA, 2000.
8. Peleg M, Tu S, Bury J, Ciccarese P, Fox J, Greenes R. Hall A, R., Johnson P. D, Jones N, Kumar A, Miksch, Quaglini S, Seyfang A, Shortliffe E H & Stefanelli M. Comparing Computer-Interpretable Guideline Models: A Case-Study Approach. *J Amer. Med. Informatics Association*, vol. 10, No. 1, Jan-Feb 2003, pp. 52-68 2002Benson and Page, 2004-08-08
9. de Clercq P et al "approaches for creating computer-interpretable guidelines that facilitate decision support" *Artificial Intelligence in Medicin*e, 2004, 31, 1-27.
10. Sutton D R, Fox J. "The Syntax and Semantics of PROforma" *J Am Med Inform Assoc*. 2003 Sep-Oct;10(5):433-43.
11. Walton R et al, "Effects of decision support on prescribing in general practice". *British Medical Journal*, 26th September. 1997
12. Jonathan Bury, Chris Hurt, Anindita Roy, Louise Cheesman, Mike Bradburn, Simon Cross, John Fox, Vaskar Saha "LISA - a web-based trial management and decision support system for childhood Acute Lymphoblastic Leukaemia" *B J Haematology*, 2005.
13. Coulson A., Glasspool D., Fox J., & Emery J. "RAGs: A novel approach to computerised genetic risk assessment and decision support from pedigrees" *Methods of Information in Medicine* 2001; 40; 315-322.
14. Emery J, Walton R, Murphy M, Austoker J, Yudkin P, Chapman C, Coulson A, Glasspool D, Fox J, "Computer support for interpreting family histories of breast and ovarian cancer in primary care: comparative study with simulated cases" *British Medical Journal*, 321, 28-32, 1st July, 2000.
15. Glasspool, D., Fox J., Castillo F D., Monaghan V. "Interactive decision support for medical planning" *Proceedings of the 9th Conference on Artificial Intelligence in Medicine in Europe* (AIME-03), 335-340, Berlin: Springer, 2003
16. Glasspool D and Oettinger, A (2005) "Interactive Decision Support for Medical Planning" Submitted for publication.
17. Tural C et al. "Clinical utility of HIV-1 genotyping and expert advice: the Havana Trial". *AIDS*. 2002; 16:209-218.
18. Taylor P, Fox J Pokropek A The development and evaluation of CADMIUM: a prototype system to assist in the interpretation of mammograms, *Medical Image Analysis*, Volume 3, Issue 4, 1999, pp. 321-337.
19. Patkar, Hurt et al "Evaluation of a computerised decision support system for the triple assessment of women referred to specialist breast cancer units" (submitted for publication, 2006).
20. Huang J, Jennings N and Fox J "Co-operation in distributed medical care" *Proc. 2nd Int'l Conference on Co-operative Information Systems*, 255-263, Toronto 1994.
21. Huang, J Jennings N R, and Fox J "An agent based approach to health care management" *Int. J. Applied Artificial Intelligence* 9 (4) 1995.
22. Huang J, Fox J, Gordon C and Jackson-Smale A "Symbolic decision support in medical care" *Artificial Intelligence in medicine*, 5, 415-430 (1993).
23. Fox J, Krause P "Arguments, contradictions and practical reasoning" *Proc. ECAI*, 1992, 623-627, Wiley.
24. Black E Using Agent Technology to Model Complex Medical Organisations *AgentLink News* issue 12, April 2003
25. Black E, Fox J "Breast Cancer Referral System with Queuing Strategies" Cancer Research UK, Advanced Computation Lab., Technical Report, 2002.