

A multicentre integration of a computer-led follow-up of prostate cancer is valid and safe

Hesham A Salem*[¶] , Giacomo Caddeo*, Jon McFarlane[†], Kunjan Patel*, Lynda Cochrane[‡], Daniele Soria[§], Mike Henley* and Jonathan Lund[¶]

*Derby Hospital NHS Foundation trust, Derby, UK, [†]Royal United Hospital Bath NHS Trust, [‡]Clinical Statistics Consultants, Dundee, UK, [§]Department of Computer Science, University of Westminster, London, UK, and [¶]Clinical Sciences Wing, The Medical School, University of Nottingham, Nottingham, UK

Communication to society

An abstract of this research was presented at The Society of Academic and Research Surgery (SARS) 2017 meeting (presentation number 089) and then published in the *British Journal of Surgery* (*Brit J Surg* 2017; 104 [Suppl. 53]).

Objective

To test a computer-led follow-up service for prostate cancer in two UK hospitals; the testing aimed to validate the computer expert system in making clinical decisions according to the individual patient's clinical need with a valid model accurately identify patients with disease recurrence or treatment failure based on their blood test and clinical picture.

Patients and Methods

A clinical-decision support system (CDSS) was developed from European (European Association of Urology) and national (National Institute for Health and Care Excellence) guidelines along with knowledge acquired from Urologists. This model was then applied in two UK hospitals to review patients after prostate cancer treatment. These patients' data ($n = 200$) were then reviewed by two independent urology consultants (blinded from the CDSS and the other consultant's rating) and the agreement was calculated by kappa statistics for validation. The second endpoint was to verify the system by estimating the system reliability.

Introduction

Prostate cancer is the most common malignant disease in men and the third leading cause of cancer-related mortality in the UK in 2014 [1]. It is a disease of the elderly, with men aged >75 years at higher risk of disease-related mortality [2]. It has a heterogeneous course of progression from slow growing and potentially insignificant to aggressive with serious impact on the patient's health and quality of life. Despite the disease's rising prevalence, there has been a significant improvement in survival rates [3,4], which may be accounted for by the availability of more accurate diagnostic and treatment modalities. However, where best to safely follow-up patients, remains unresolved. The current National

Results

The two individual urology consultants identified 12% and 15% of the patients to have potential disease progression and recommended their referral to urology care. The kappa coefficient for the agreement between the CDSS and the two consultants was 0.81 ($P < 0.001$) and 0.84 ($P < 0.001$). The agreement amongst both specialist was also high with $k = 0.83$ ($P < 0.001$). The system reliability was estimated on all cases and this demonstrated 100% repeatability of the decisions.

Conclusion

A CDSS follow-up is a valid model for providing safe follow-up for prostate cancer.

Keywords

clinical decision support, expert system, prostate cancer follow-up, knowledge validation, rule-based systems, system validation and verification, #PCSM, #ProstateCancer

Institute for Health and Care Excellence (NICE) recommendation is to provide this follow-up in primary care where appropriate [5]. However, this recommendation was met with concerns from both GPs and Urologists because of the lack of expertise in the community [6]. The National Prostate Cancer Audit 2014 annual report identified five different models for prostate cancer follow-up: Consultant-led clinic, Cancer Nurse Specialist-led clinic, Telephone clinic, Community-based specialist follow-up, and Radiographer-led clinic (only for radiotherapy pathway) [7]. The community-based model played only a minor role in follow-up in this audit.

In various industries expert-simulating systems have provided an alternative cheap, reliable and available solution where

expertise is lacking [8]. The bottleneck to their development is the knowledge acquisition phase, which usually is exhaustive and time consuming [9]. Furthermore, the system has to go through rigorous testing; system validation (are we building the right system?) and system verification (are we building the system correctly?). Previous attempts to apply this to an industry setting have suffered from a lack of consistency and formality in the model [10]. The same problems were also evident in medical expert system development and has challenged their uptake in healthcare [11].

In the present study, we aimed to verify and validate an expert system simulating a Urologist in the follow-up of stable prostate cancer and its application in supporting GPs in providing a follow-up service in the community. Therefore, we tested the null hypothesis that the agreement between the expert system and the human domain experts is slight to fair ($k_0 = 0.4$) against the alternative that it is better than moderate ($k_1 > 0.6$). For verification, the expert system has to be 100% reliable, i.e. the same clinical input triggered the system to produce the exact same outputs on each occasion.

Patients and Methods

Description of the Prostate Cancer System

The present clinical-decision support system (CDSS) is a web-based solution that can be used with current versions of browsers including Microsoft IE, Google Chrome, Mozilla, and Safari. At the core of the system is a rule engine, which processes all of the data captured and held for a patient, and presents it to a complex disease domain-specific algorithm compiled and tested in accordance with NICE guidelines [5]. The system uses principally Microsoft technologies including: Microsoft SQL Server database, the Microsoft.NET development framework and the C# programming language, along with other web technologies, e.g. JavaScript and AJAX (Asynchronous JavaScript And XML).

System security and Functional Analysis

The expert system is based on a secure remote server held within the NHS Information Technology (IT) network (N3). This location makes it accessible to all NHS primary and secondary care sites, and NHS partners without the need for individual copies at each site, providing one single point of access. Only authorised personnel can gain access to this system and all the stored data are encrypted consistent with the standards of the Information Governance Statement of Compliance (IG SoC).

The system is triggered by an authorised-user query about a patient visit. The CDSS accesses their records on the Hospital Information System, including their laboratory test results,

before asking the user to complete a problem-specific online form. The form includes routine questions enquiring about relevant symptoms as outlined in the NICE 2014 guidelines to support the system in accurate decision-making [5]. The system output is in the form of an electronic clinic letter suggesting a tailored treatment plan for the individual patient consultation.

Study Configuration

The CDSS is already part of an existing cancer nurse specialist-led prostate cancer follow-up service in Derby Hospital NHS Foundation trust and University Hospitals Coventry and Warwickshire NHS Trust. The main system function in this service is acting as an intelligent database and clinic letter generator (Figs 1 and 2). Patients included were those who had their disease diagnosed and treated by the urology cancer multidisciplinary team before being discharged for follow-up to the cancer nurse specialist clinic. There were no specific criteria for follow-up in this clinic and the clinic predated the software.

When seeing a patient for follow-up the users (cancer nurse specialist in this study) supply the CDSS with recent symptoms and blood test results, and the system produces an electronic letter stating significant clinical details and a suggested treatment plan. The treatment plan was reviewed and overridden by the cancer nurse specialist if appropriate to do so.

In this study, we only aimed to evaluate the systems' validity, thus the letters were anonymised and the nurses' plans were eliminated before examining the cases against the inclusion criteria. Each letter generated by the CDSS had five main components (Figs 1 and 2):

- 1 History of disease and treatment.
- 2 Current symptoms and blood test results.
- 3 System analysis of clinical data.
- 4 System suggested outcome.
- 5 Nurse notes text box for supplying any extra information and overriding the system decision where necessary.

The eligible cases were anonymised by eliminating all patients' identifiers and the cancer nurse specialist free-text entry of their own clinical assessment and plan (component 5 in Fig. 2).

Cases were examined by the study clinical monitor against the inclusion criteria and cases were excluded from the study with justification (as not enough clinical data, newly diagnosed or treated cases i.e. stable not prostate cancer etc.).

Eligibility Criteria

All adult patients seen in this clinic with known prostate cancer whom had their primary treatment assigned and delivered.

Fig. 1 A sample of the anonymised electronic letter generated by the clinical-decision support system in the live clinic (page one).

Department of Urology
 Prostate Cancer Telephone Clinic
 NHS Number:
 Date of Clinic:
 Urology:
 Clinical oncology:

PRIVATE AND CONFIDENTIAL

Direct Line:

Dear Dr, [redacted] Follow on page contains blood test results and nurse's notes.

[redacted]

1 Patient Summary
 Stable Prostate Cancer Pathway : External Beam Radiation Therapy (EBRT)

Mr [redacted] was reviewed today in the stable prostate cancer clinic. Mr [redacted] is 59 years old and was diagnosed with gleason score 3+4 cancer on 25 October 2006. His TNM stage was recorded as pT3 N0 M0.

His recorded PSA at presentation was 8.40 ng/ml. We have no record of a formal PSA nadir, we have therefore calculated a nadir (0.04 ng/ml) based on our available PSA values.

2 Treatments:
 24 Nov 2006 Zometa trial

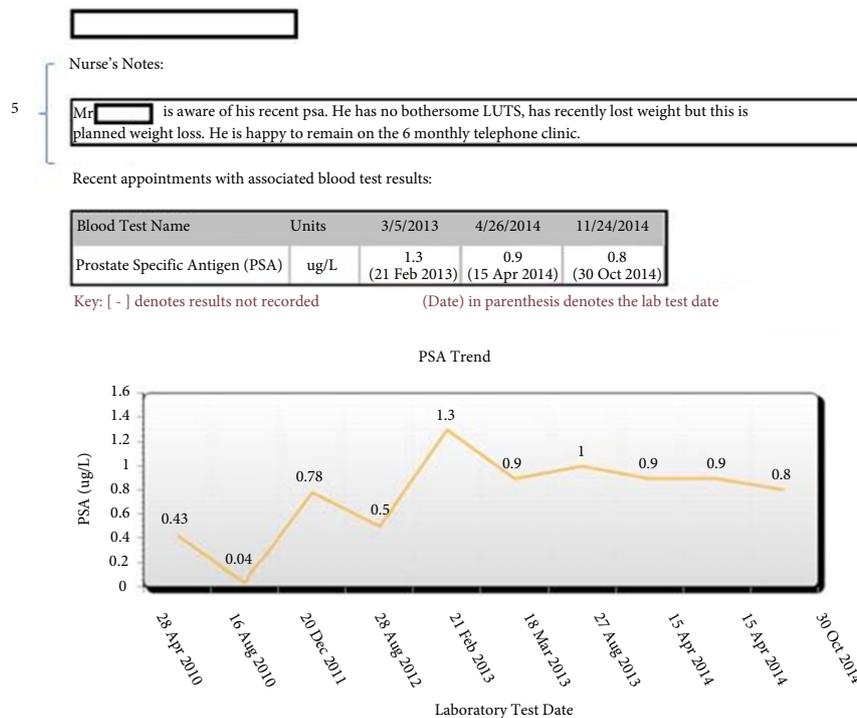
Today's Findings
 The level of PSA present in the last available blood test is 0.80 ng/ml.

3 Clinical Management
 The patient is being monitored on the radiotherapy management pathway, his PSA is within acceptable parameters.
 We are using a calculated nadir (0.04 ng/ml) to make our assessments as a formal hospital PSA nadir is unavailable.
 We do not have results for potassium, creatinine, alkaline phosphatase and calcium.

4 Follow Up
 An appointment will be made to see the patient again in 6 months time.

Kind Regards,
 Nurse
 On behalf of the stable prostate cancer clinic

Fig. 2 A sample of the anonymised electronic letter generated by the clinical-decision support system in the live clinic (page two).



Inclusion Criteria

Adult patients with known prostate cancer and presenting for routine follow up in clinic with (essential requirement):

- 1 Known prostate cancer treatment pathway.
- 2 Details on his serum PSA levels on presentation and most recent results.

With or without (optional requirement):

- 1 Disease or treatment-related symptoms (lower urinary tract, bone pain, weight loss and erectile dysfunction).
- 2 Abnormal blood test (haemoglobin, calcium, urea, electrolytes, creatinine, and liver function tests).

Exclusion Criteria

- 1 Cases with newly diagnosed prostate cancer that are waiting for a decision to treat.
- 2 Cases with known prostate cancer that are being evaluated for known recent disease progression.
- 3 Cases with no clinical details of their initial cancer presentation such as Gleason score, PSA level, and tumour grade, which would influence the follow-up decision-making.

In order to reach both endpoints, the study divided into two branches after case selection (Fig. 3). In the primary endpoint pathway, the included cases were reviewed by two independent clinical investigators after eliminating the system analysis and suggestions (Components 3 and 4 samples in Fig. 1). Each investigator independently assigned an outcome for each case (Fig. 4) according to the disease history, current results, and symptoms blinded from the system recommendation.

In the second part of the study, the system reliability and precision was estimated by codifying all cases clinical variables and output. The cases specific codes were all tabulated on a spread sheet (Excel, Microsoft Corp, Seattle, WA, USA) and all cases with similar input(s) were expected to have the same output code. In case of discrepancy, errors were identified and the system was then corrected. Any system changes were followed by a re-test by a new sample of cases to re-estimate reliability until 100% precision was obtained (secondary endpoint).

Statistical Evaluation

The system validation was estimated by testing the null hypothesis that the agreement between the CDSS and the human domain experts was slight to fair ($k_0 = 0.4$) against the alternative that it was better than moderate ($k_1 > 0.6$) [12]. Kappa was estimated between each investigator and the CDSS then between both investigators.

Results

The study included data of 200 patients seen in either hospital (100 each; Fig. 5). All patients had prostate cancer, and the mean and median age was 75 (range 51–94) years, with a mean and mode Gleason score of 7. The risk stratifications according to D'Amico's classification [13] of their disease on diagnosis identified 96 (48%) patients in the high-risk group, and 78 (39%) and 26 (13%) in the intermediate- and low-risk groups, respectively (Fig. 6). Radiotherapy had been used to treat 128 cases out of the total (Fig. 7).

Kappa statistics were estimated to test the hypothesis. The unweighted kappa for investigator 1 and 2 was 0.8, and with quadratic weighting the kappa remained, demonstrating substantial agreement between the CDSS and investigator at 0.86 and 0.96 for investigator 1 and investigator 2, respectively. This refuted the null hypothesis as $k_0 > 0.4$ and alternative hypothesis $k_1 > 0.6$ was accepted, with the conclusion that the CDSS is valid and its assigned outcomes are acceptable by the domain experts (Tables 1, 2 and 3).

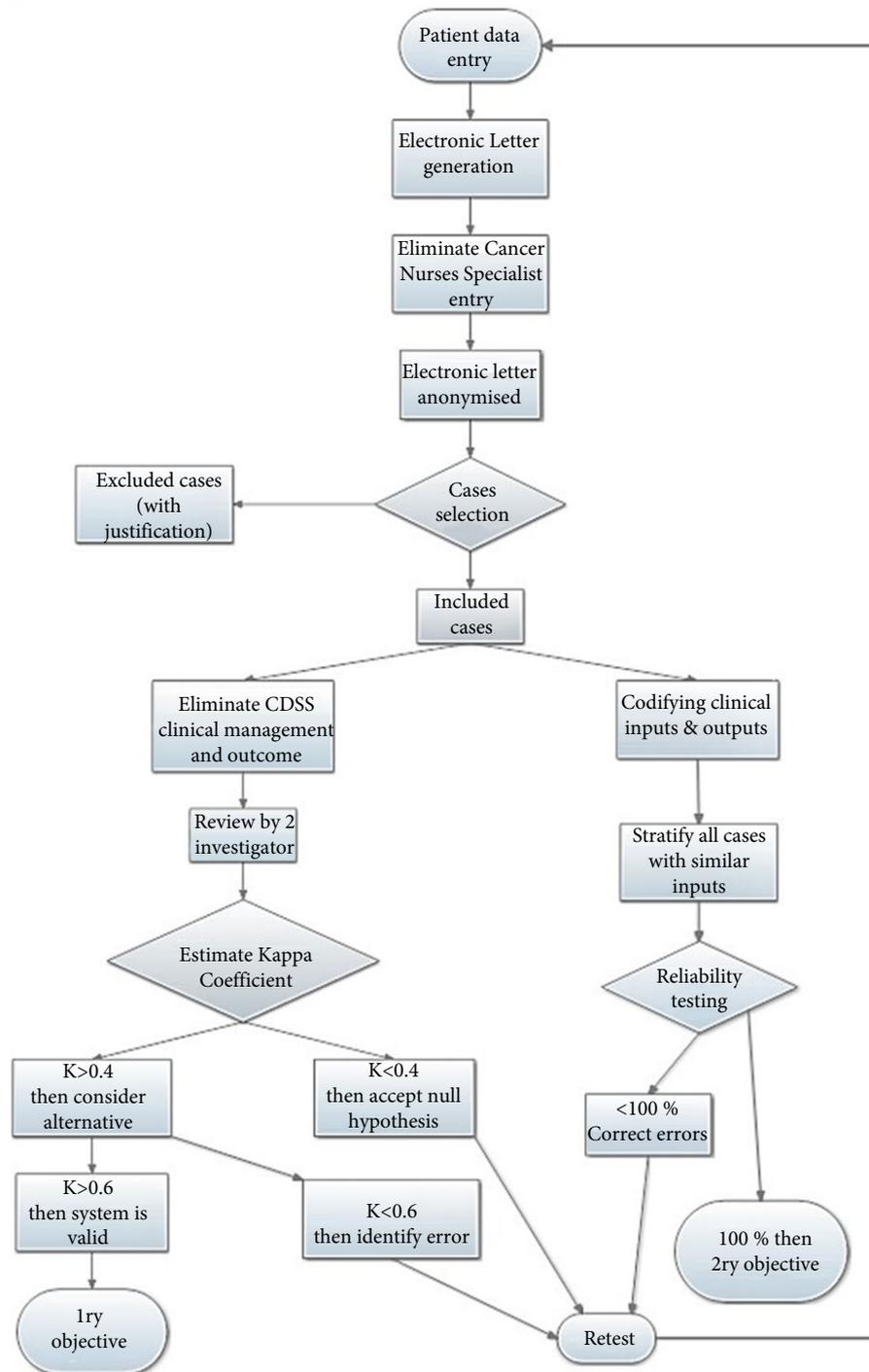
Both investigators identified 31 cases with possible disease progression or recurrence (investigator 1, 25; investigator 2, 29). Most of those cases were biochemical failure after radiotherapy treatment (21 out of the total of 31) based on either American Society for Therapeutic Radiology and Oncology (ASTRO) [14–16] or Phoenix [17] criteria, or both. Six cases were on watchful waiting and had either high or rising PSA levels and were identified as disease progression (those patients did not receive any treatment). Two patients had radical surgery with detectable PSA levels and one was treated with hormone ablation, with a significant PSA-doubling time. The CDSS identified all of the above cases as disease progression or recurrence except one case who was only classified by investigator 1 as potential recurrence after radiotherapy. They had only two consecutive rises in their PSA level, thus not identified as biochemical failure by the CDSS or investigator 2.

The reliability testing was estimated by examining the outcome of cases with similar input codes. This showed 100% reliability and so the study secondary objective was met.

Discussion

The present multicentre study validated and verified a novel CDSS-led follow-up service that can be used in a prostate cancer follow-up setting, which could be based in either primary or secondary care. The present study is the first to combine quantitative methodologies to perform full system validation and verification on real patients. This critical test has been lacking in the development of medical expert systems, and this has led to poor uptake because of a lack of confidence in these systems [11]. Standalone software as

Fig. 3 Live validation study flow chart displaying both endpoints of the study.



CDSS is recognised by European regulatory bodies as a medical device and require registration via strict criteria [18]. These criteria are outlined in the declarations of the International Organization for Standardization (ISO) 13 485 (quality management systems), 14 971 (risk management) and 62 304 (software development cycle). The latter addresses the standards for testing that should be met in the validation.

In the present study, we quantified the validity by estimating the kappa statistics in a two-centre live clinical evaluation. Furthermore, the objective reliability testing verified the system and completed the development cycle. These quantified tests support the utility of the expert system, as it has met state of the art verification and validation methodologies [18].

Fig. 4 A sample of the clinical data form. Those were anonymised and had their clinical outcome omitted before being presented to the individual investigator for rating.

Department of Urology
 Prostate Cancer Telephone Clinic
 NHS Number: 0123456789
 Date of Clinic: 26 Jan 2015
 Consultant: MJH
 Study case DB268

Re: Mr Study case DB268

Patient Summary

Stable Prostate Cancer Pathway: *External Beam Radiation Therapy (EBRT)*

Mr DB268 was received today in the stable prostate cancer clinic. Mr DB268 is 74 years old and was diagnosed with gleason score 6 cancer on 12 May 2004. His TNM stage was recorded as pT2cN0 M0.

His recorded PSA at presentation was 7.90 ng/ml. His PSA nadir is [0.08 ng/ml].

Treatments:

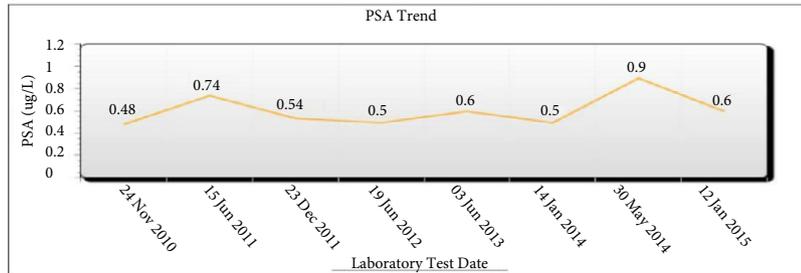
15 Nov 2004 Radical radiotherapy with neo-adjuvant hormonal therapy completed December 2004.

Today's Findings

The level of PSA present in the last available blood test is 0.60 ng/ml. no other symptoms.

Clinical Management

Blood test Name	Units	6/17/2013	1/30/2014	6/16/2014	1/26/2015
Creatinine	umol/L	-	92 (14 Jan 2014)	-	-
Potassium	mmol/L	-	4 (14 Jan 2014)	-	-
Prostate specific antigen (PSA)	ug/L	0.6 (03 Jun 2013)	0.5 (14 Jan 2014)	0.9 (30 May 2014)	0.6 (12 Jan 2015)



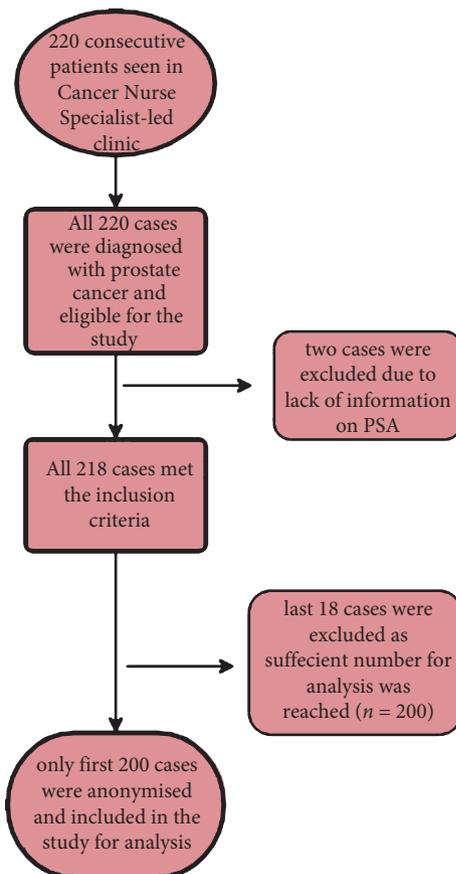
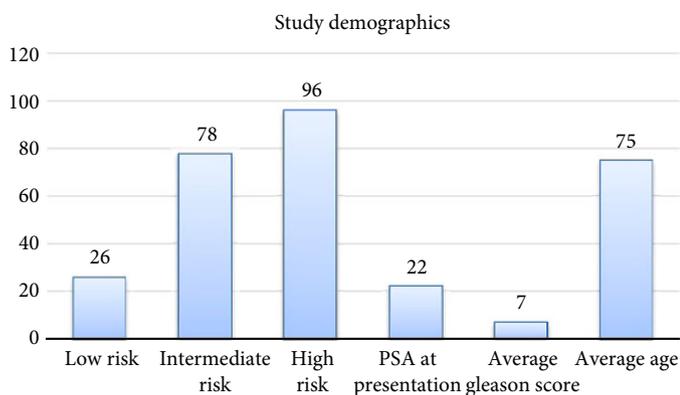
OUTCOME	Comments
1	
2	
3	
4	

1 = ROUTINE, 2 = ROUTINE WITH ADVICE, 3 = URGENT, 4 = REFER TO HOSPITAL.

NAME	INVESTIGATOR
DATE	
SIGN	

It has been estimated that prostate cancer costs more than €8.43 billion across Europe (£0.8 billion in the UK), with most spent on treatment in the first year after diagnosis [19].

Prostate cancer follow-up is also expensive overall and cost-effective service improvement is much needed. This cost is expected to increase with improvement in disease-specific

Fig. 5 Included cases flow chart.**Fig. 6** Included cases demographics. Cases were stratified according to their data in the initial diagnosis of prostate cancer following D'Amico's classification criteria. Most were in the intermediate- and high-risk group.

survival and increase in the population life-expectancy. The CDSS model may be able to safely move follow-up in to primary care with potentially significant cost saving [6]. The system developers believe that the valid system can be used by any healthcare worker regardless of their urological background, with a potential of adopting an interface for

direct patient interaction. This could have significant benefits and cost savings but given the age group of the patients and lack of flexibility of NHS healthcare, the later would be challenging to implement.

This type of study is limited by the quality of the clinical data and experience of the human experts. The data in our present study were all real patients' data collected in real-time clinics from two large tertiary and secondary care centres. Furthermore, uncommon and rare presentations have been validated in the knowledge validation study. Cases studied were allocated to the cancer nurse specialist follow-up clinic by the multidisciplinary team and may explain the low recurrence rate in the present study, even though 48% of the cases were in the high-risk group.

The clinical investigators were both qualified consultants and they received independent specialist training. One had special interest in prostate cancer community follow-up and had previously developed a follow-up model for a remote area. The other investigator had an interest in core urology and stone disease. Both were not familiar with the software development and had no conflict of interest. The rating was performed blinded from the other investigator and the CDSS, thus bias was eliminated and the agreement calculated by the kappa coefficient should be the true agreement.

In medicine, the validation of an intervention usually requires a comparison against the 'gold standard' to estimate the sensitivity and specificity. This is different to expert system validation, where it is based on the Turing test concept in comparing the machine's cognitive performance to human domain experts [10,20]. Kappa statistics are more appropriate for this type of validation, as they estimate the agreement between human and machine cognitive function and have been applied by other studies to validate expert systems [21,22]. The sensitivity and specificity is only accurate if compared with the 'gold standard' of histopathology results and this is more widely applied by artificial neural networks, as they learn directly from data and the role of human opinion is limited [9].

The long-term outcome of the cases was not examined as part of the present study. The validation aimed to estimate the validity of the system reasoning against consultant urologists. The long-term outcome can be useful to estimate the accuracy of both investigators and the CDSS; however, this would require long-term follow-up and more invasive tests.

Conclusion

The CDSS demonstrated high validity and accuracy in its decision making. This tool has the potential for safe use in supporting follow-up of surgical oncology in primary and

Fig. 7 Cases stratification according to treatment. The stratification was according to the most significant treatment on which we calculate biochemical failure and disease progression. Radiotherapy was the most common, as this group usually requires longer follow-up than patients after surgical treatment.

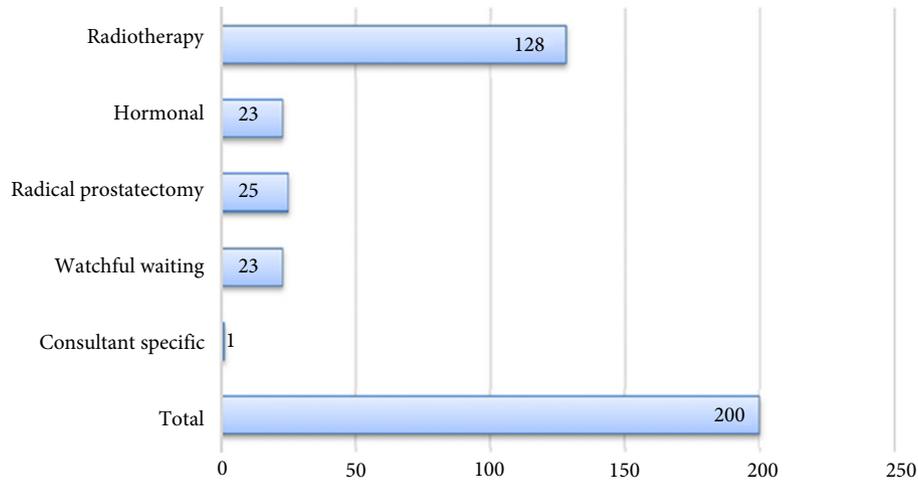


Table 1 Kappa statistics for investigator 1 and CDSS.

Type of kappa	Observed kappa	Standard error	95% CI
Unweighted	0.8	0.043	0.72–0.89
Linear weighting	0.8	0.04	0.75–0.91
Quadratic weighting	0.86	0.046	0.78–0.94

Table 2 Kappa statistics for investigator 2 and CDSS.

Type of kappa	Observed kappa	Standard error	95% CI
Unweighted	0.85	0.039	0.77–0.92
Linear weighting	0.92	0.023	0.86–0.96
Quadratic weighting	0.96	NC*	NC*

*N, this quantity cannot be calculated.

Table 3 Kappa statistics for investigators 1 and 2.

Type of kappa	Observed kappa	Standard error	95% CI
Unweighted	0.83	0.43	0.74–0.91
Linear weighting	0.85	0.4	0.77–0.93
Quadratic weighting	0.87	0.043	0.79–0.95

secondary care. A mixed methodology approach is required to perform the mandatory system validation and verification.

Funding

We have received no funding to complete this research.

Conflict of Interest

None.

References

- 1 Cancer Research UK. Prostate cancer mortality 2016. Available at: <http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/prostate-cancer#heading-One>. Accessed February 2018.
- 2 Droz JP, Balducci L, Bolla M et al. Management of prostate cancer in older men: recommendations of a working group of the International Society of Geriatric Oncology. *BJU Int* 2010; 106: 462–9
- 3 Cancer Research UK. Prostate cancer survival statistics, 2014. Available at: <http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/prostate-cancer/survival>. Accessed February 2018.
- 4 National Cancer Institute, Surveillance, Epidemiology and End Results (SEER) program. Cancer Stat Facts: Prostate Cancer, 2014. Available at: <https://seer.cancer.gov/statfacts/html/prost.html>. Accessed February 2018.
- 5 The National Institute for Health and Care Excellence (NICE). Prostate cancer: diagnosis and management. NICE guidelines [CG175], 2014. [Online]. Available at: <https://www.nice.org.uk/guidance/cg175>. Accessed February 2018.
- 6 Watson EK, O'Brien R, Campbell C et al. Views of health professionals on the role of primary care in the follow-up of men with prostate cancer. *Fam Pract* 2011; 28: 647–54
- 7 National Prostate Cancer Audit (NCPA). NCPA 2014 Annual Report. Available at: <http://www.npca.org.uk/reports/>. Accessed February 2018.
- 8 Liao SH. Expert system methodologies and applications - a decade review from 1995 to 2004. *Expert Syst Appl* 2005; 28: 93–103
- 9 Pandey B, Mishra RB. Knowledge and intelligent computing system in medicine. *Comput Biol Med* 2009; 39: 215–30
- 10 Okeefe RM, O'leary DE. Expert system verification and validation - a survey and tutorial. *Artif Intell Rev* 1993; 7: 3–42
- 11 Ammenwerth E, Nykanen P, Rigby M, de Keizer N. Clinical decision support systems: need for evidence, need for evaluation. *Artif Intell Med* 2013; 59: 1–3
- 12 Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960; 20: 37–46
- 13 D'Amico AV, Whittington R, Malkowicz SB et al. Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. *JAMA* 1998; 280: 969–74

- 14 Bellera CA, Hanley JA, Joseph L, Albertsen PC. A statistical evaluation of rules for biochemical failure after radiotherapy in men treated for prostate cancer. *Int J Radiat Oncol Biol Phys* 2009; 75: 1357–63
- 15 Kuban DA, Levy LB, Potters L et al. Comparison of biochemical failure definitions for permanent prostate brachytherapy. *Int J Radiat Oncol Biol Phys* 2006; 65: 1487–93
- 16 Horwitz EM, Thames HD, Kuban DA et al. Definitions of biochemical failure that best predict clinical failure in patients with prostate cancer treated with external beam radiation alone: a multi-institutional pooled analysis. *J Urol* 2005; 173: 797–802
- 17 American Society for Therapeutic Radiology and Oncology Consensus Panel. Consensus statement: guidelines for PSA following radiation therapy. American Society for Therapeutic Radiology and Oncology Consensus Panel. *Int J Radiat Oncol Biol Phys* 1997; 37: 1035–41
- 18 The European, Parliament the Council of the European Union. 93/42/EEC. Directive 2007/47/EC of the European Parliament and of the Council of 5 September 2007. Available at: http://ec.europa.eu/consumers/sectors/medical-devices/files/revision_docs/2007-47-en_en.pdf. Accessed February 2018.
- 19 Luengo-Fernandez R, Leal J, Gray A, Sullivan R. Economic burden of cancer across the European Union: a population-based cost analysis. *Lancet Oncol* 2013; 14: 1165–74
- 20 Turing AM. Computing machinery and intelligence. *Mind* 1950; 49: 433–60
- 21 Lopes MH, Ortega NR, Silveira PS, Massad E, Higa R, Marin Hde D. Fuzzy cognitive map in differential diagnosis of alterations in urinary elimination: a nursing approach. *Int J Med Inform* 2013; 82: 201–8
- 22 Torshizi AD, Zarandi MH, Torshizi GD, Eghbali K. A hybrid fuzzy-ontology based intelligent system to determine level of severity and treatment recommendation for benign prostatic hyperplasia. *Comput Methods Programs Biomed* 2014; 113: 301–13

Correspondence: Hesham A Salem, PhD Student, University of Nottingham, The Medical School, Royal Derby Hospital, Uttoxeter Road, Derby DE22 3DT, UK.

e-mail: hesham.salem1@nhs.net

Abbreviations: CDSS, clinical-decision support system; NICE, National Institute for Health and Care Excellence.