

Expert systems development and the validation of computer led follow up in prostate cancer

By

Hesham Salem MB Bch, MRCS, MRCS DOHNS

Thesis submitted to The University of Nottingham for the Degree of Doctor of Philosophy, (PhD Nottingham)

March 2017

School of Medicine and Health Care

The University of Nottingham

Nottingham, United Kingdom

DEDICATED WITH LOVE TO MY PARENTS, MY CHILDREN AND ALL MY FAMILY MEMBERS …

Acknowledgment

First and above all I praise almighty Allah for granting me the opportunity to learn and complete this research. It would not have been possible without the support of Dr Lorraine Pinnington the director of postgraduate education, Karen Kirkland and all school office staff at the Division of Medical Sciences.

I would like to acknowledge the great effort of Mr Mike Henley and Mr Steven Pick for developing the prototype of the Expert system and Dr Sue Hagan for planning the study for the first clinical investigation of this prototype. I am grateful for Mr Mike Henley's and Mr Steven Pick's effort in developing the full version of the software and adopting my validation and verification studies of their software. I would like to extend my appreciation to Mr Kunjan Patel for his effort in managing and anonymising the database.

I would like to thank Professor Ronald Summers for his support and suggestions for formulating the methods for my thesis. I would like to extent my appreciation to Mr Mohamad Darweesh for explaining the verification and validation process and his valuable suggestion to their integration in the system testing. This research gained valuable knowledge from the clinical investigators in the knowledge acquisition study and for this I would to thank Royal Derby Hospital Urologist, Mr Stephen Thomas, Mr Amjad Peracha, Mr Jas Ray, Mr F Amawi, Mr Simon Williams and Mr Huw Williams. I would like to express my sincere appreciation to the fine work and dedication provided by the Mr Giacomo Caddeo, Mr Sadmeet Singh, Richard Stanford, and John McFarlane for their remarkable effort in the validation and verification study.

I would like to thank Dr Apostolos Fakis for providing statistical support for the prototype clinical investigation and ethics approval.

I would like to express my appreciation to Dr Lynda Cochran for her great help in designing the statistical methods for the knowledge validation then the system verification. I would like to thank my PhD colleague Dr Fahd Aljebab for his valuable advice in designing my systematic review methodology.

I would like to thank my supervisor Dr Soria for his guidance, support and great effort in simplifying complex computer science for me.

Last but not least I would like to thank my clinical supervisor Mr Lund for his great help and valuable guidance. He has been very generous with his time and effort for completing this work and made such a long journey in academia enjoyable and worthwhile.

Publications and presentations related to this thesis

Publications in peer reviewed Journals

- ***Published article:*** *A multicentre integration of a computer led follow up of prostate cancer; is valid and safe'.* **Salem H**, Mr K Patel, Dr L Cochrane, Mr J Caddeo, Mr J McFarlane, Dr D Soria, Mr M Henley, Mr J Lund. British Journal of Urology International, Volume122, Issue3, Pages 418-426, September 2018
- ***Published abstract:*** *'Live validation of an experimental expert simulating model in managing prostate cancer in the community'.* **Salem H,** G Caddeo, K Patel, M Henley, D Soria, J Lund. British Journal of Surgery, Volume 103, Issue Supplement S3, pages 6–49, April 2016
- ***Published abstract:*** *Evaluation of a novel computer based model for diagnosing disease progression in patients with stable prostate cancer against gold standard decisions set by a Urology consultant.* **Salem H,** Mr R Stanford, Dr D Soria, Mr S Thomas, Mr J Lund, Mr M Henley. World Journal of Urology, October 2015, Volume 33, Supplement 1, pp 1-256
- ***Published abstract:*** *History and shortcomings of artificial intelligence in simulating Urologists.* **Salem H**, Dr D Soria, Mr J Lund. World Journal of Urology, October 2015, Volume 33, Supplement 1, pp 1-256

Oral Presentations

- *A pilot trial of a computer based cancer follow up in two teaching and university hospitals; is safe and valid.* **Salem H**, K Patel, G Caddeo, J McFarlane, M Henley, D Soria, J Lund. SARS Annual Meeting 2016, Royal College of Surgeons of Ireland.

- *'Live validation of an experimental expert simulating model in managing prostate cancer in the community'* **Salem H,** K Patel, G Caddeo, M Henley, D Soria, J Lund SARS Annual Meeting 2016, Royal College of Surgeons of England.
- *A medical auto pilot system for cancer patients in Derby?* **Salem H**, PhD student in Medicine and Health care at The University of Nottingham. The University of Nottingham Research Show case June 2014. (Won Third place for best presentation).

Poster presentation

- *Evaluation of a novel computer based model for diagnosing disease progression in patients with stable prostate cancer against gold standard decisions set by a Urology consultant.* **Salem H**, Mr R Stanford, Dr D Soria, Mr S Thomas, Mr J Lund, Mr M Henley. **(Moderated poster)** 35th Congress of the Societe Internationale d'Urologie 2016, Melbourne, Australia.
- *History and shortcomings of artificial intelligence in simulating Urologists* **Salem H**, Dr D Soria, Mr J Lund. **(Unmoderated poster)**. 35th Congress of the Societe Internationale d'Urologie 2016, Melbourne, Australia.
- *'Virtual Urology clinic' a feasibility study in a busy UK teaching Hospital.* **(Unmoderated poster) Salem H**, Mr S Thomas. 35th Congress of the Societe Internationale d'Urologie 2016, Melbourne Australia.
- *Use of a Computer Decision Expert System to support the follow up of stable prostate cancer in primary care is valid and safe* **(poster) Salem H**, Dr D Soria, Mr J Lund, Mr M Henley. Royal College of General Practitioners Annual Conference Oct 2015.

- *Artificial Intelligence in Urology, their strength, shortfalls and state of the art* **(poster) Salem H**, Dr D Soria, Mr J Lund. The 11th World Congress on Urological Research (ESUR-SBUR15), EAU, Nijmegen, Netherland, Sep 2015.
- *A prospective validation of a Computer based follow up for patients with prostate cancer.* **(poster) Salem H**, Mr R Stanford, Mr S Singh, Dr D Soria, Mr J Lund, Mr M Henley. The 11th World Congress on Urological Research (ESUR-SBUR15), EAU, Nijmegen, Netherland, Sep 2015.
- *Registering expert software as a Uro-oncology specialist* **(poster). Salem H**, Awwad A, Thomas S, Peracha A, Amawi F, Henley M. The Medical Physics and Engineering Conference 2014 (MPEC2014) was held at the Scottish Exhibition and Conference Centre in Glasgow, Scotland.
- *A computer doctor managing patients with prostate cancer in the community and two UK university hospitals.* **(poster)** Falah Amawi, **Salem H**; Stephen Thomas; Amjad Peracha; Simon Williams; Steven Pike; Sue Hagen; Mike Henley. Royal College of General Practitioner annual conference, 2013, Harrogate, UK.

Table of Contents

Acknowledgment 2
Publications and presentations related to this thesis 4
Table of Contents 7
Table of Figures 11
List of Tables 14
Abbreviations 15
List of appendices 20
1 Chapter one - Introduction 21
1.1 Background 22
1.2 Expert system structure, development and validation 23
1.2.1 Expert system structure 25
1.2.2 Expert systems development life cycle: 41
1.3 Discussion 53
1.4 Summary 54
1.5 Aims and objectives of this thesis 54
1.6 Contribution to knowledge 55
1.7 Thesis organisation 55
2 Chapter two - Systematic review of Artificial intelligence in Urology 57
2.1 Introduction 58
2.2 Methods 58
2.3 Results 62
2.3.1 Quantitative analysis 64
2.3.2 Qualitative analysis 107
2.4 Discussion 112
2.5 Conclusion 114
3 Chapter three - Prostate cancer knowledge acquisition (declarative knowledge) 115
3.1 Synopsis 116
3.1.1 Background 116
3.1.2 Objectives 116
3.1.3 Results 116
3.2 Introduction 118
3.3 Risk factors 118
3.4 Prostate cancer mortality 120
3.5 Treatment pathways 121
3.5.1 Watchful waiting 122
3.5.2 Active surveillance 122

3.5.3 Hormonal ablation therapy and chemotherapy 125
3.5.4 Brachytherapy 128
3.5.5 Radical radiotherapy 129
3.5.6 Radical prostatectomy 131
3.5.7 Focal treatment 133
3.6 Post treatment follow up 133
3.6.1 Biochemical Failure definitions 134
3.7 Service organisation of follow up 136
3.8 Discussion 136
3.9 Conclusion 138
4 Chapter four - Prostate cancer knowledge Acquisition (procedural knowledge) 139
4.1 Synopsis 140
4.1.1 Background 140
4.1.2 Objectives 140
4.1.3 Results 140
4.2 Introduction 141
4.3 Previous clinical investigation of the prototype 141
4.4 Outcome of the clinical investigation 144
4.5 Methods 146
4.5.1 Cases inclusion 148
4.5.2 Cases preparation 149
4.5.3 Case rating 149
4.5.4 Sub analysis methodology 151
4.5.5 Statistical analysis 151
4.6 Results 154
4.7 Discussion 156
4.8 Conclusion 157
5 Chapter five - Knowledge validation 159
5.1 Synopsis 160
5.1.1 Background 160
5.1.2 Methods 160
5.1.3 Results 160
5.2 Introduction 161
5.3 Methods 161
5.3.1 Knowledge design 161
5.3.2 Case selection, synthesis and analysis 161
5.3.3 Study analysis 165

5.4 Results 167
5.4.1 Testing set 167
5.4.2 Validation set 168
5.4.3 Qualitative analysis 169
5.4.4 Quantitative analysis 169
5.5 Discussion 170
5.6 Conclusion 172
6 Chapter six - prospective multicentre validation of the prostate cancer follow up expert system 173
6.1 Synopsis 174
6.1.1 Background 174
6.1.2 Methods 174
6.1.3 Results 174
6.2 Introduction 176
6.3 Methods 176
6.3.1 Description of the prostate cancer system 176
6.3.2 System security and functional analysis 176
6.3.3 Study configuration (protocol in appendix 8) 177
6.3.4 Statistical evaluation 180
6.4 Results 182
6.5 Discussion 185
6.6 Conclusion 187
7 Thesis Conclusion 188
7.1 Summary 188
7.2 Practice implication 189
7.3 Limitation 190
7.4 Future 190
7.5 Conclusion 191
8 References 192
9 Appendices 214
Appendix 1: Protocol of the clinical investigation plan 214
Appendix 2: Refined decision rules of the prototype 275
Appendix 3: Service evaluation approval 277
Appendix 4: Protocol of the knowledge validation study 284
Appendix 5: Clinical cases presented to investigators for rating in the knowledge validation study 309
Appendix 6: Electronic letter generated by the CDSS in the live validation study 311

Appendix 7: Example of cases generated for the live validation 313
Appendix 8: Protocol of the live validation 314

Table of Figures

Figure 1.1 Expert system components 25

Figure 1.2 Example of a rule from MYCIN 28

Figure 1.3 Example of Forward chaining 29

Figure 1.4 Backward chaining 30

Figure 1.5 Artificial neural network structure 32

Figure 1.6 Single perceptron structure 33

Figure 1.7 Perceptron input nodes weighting to trigger outcome 33

Figure 1.8: An example of a single layer perceptron machine learning on cases with known prostate cancer 35

Figure 1.9 Fuzzy Logic modelling 37

Figure 1.10 Fuzzy rule base 38

Figure 1.11 Fuzzy neural modelling 39

Figure 1.12 An example of fuzzy neural modelling calculations 40

Figure 1.13 Information system development cycle 41

Figure 1.14 Types of knowledge 44

Figure 1.15 Rule based reasoning 44

Figure 1.16 Knowledge acquisition steps and objectives 46

Figure 1.17 Knauf et al. expert system five stage system validation cycle 50

Figure 1.18 Expert system verification models 52

Figure 1.19 Expert system evaluation 53

Figure 2.1 PRISMA flow chart for the systematic review of articles included in the review of expert systems in Urology 62

Figure 2.2 Analysis of Expert systems by models. 63

Figure 2.3 Expert system analysis by year of publication demonstrating an upward trend and increase in number of publications. 63
Figure 2.4 Urological domains applied by Expert systems. Prostate cancer (CaP) was the commonest domain followed by bladder cancer (Bca) then other diseases as benign prostatic disease (BPD), pelvi ureteric junction obstruction (PUJ), Urinary tract infection (UTI), renal cell cancer (RCC), vesico ureteric reflux (VU reflux). 64
Figure 3.1 European Age Standardised incidence rate of prostate cancer in UK years 118
Figure 3.2 Prostate Cancer, Average Number of New Cases per Year & Age-Specific Incidence Rates, UK 119
Figure 3.3 Number of deaths from cancer and age-specific rates of deaths from cancer (per 100,000) in England, 2016. Positive correlation of prostate cancer disease specific mortality and age..... 120
Figure 3.4 Directly age-standardised rates per 100,000 people of newly diagnosed cases of prostate cancer and cancer related mortality, England, 2006 to 2016, Source: National Cancer Registration and Analysis Service within Public Health England; Office for National Statistics 121
Figure 4.1 Flow chart for the clinical investigation plan 143
Figure 4.2 Mind map for coding the cases input and outcome 146
Figure 4.3 Decision Rule example 147
Figure 4.4 A sample of the cases before rating by the investigators 150
Figure 4.5 Study flow chart 153
Figure 5.1 Decision trees development and validation 166
Figure 5.2 Cases classified in groups according to their clinical presentation 167
Figure 5.3 Included cases stratification according to the treatment pathway 168

Figure 6.1 Live validation study flow chart 181

Figure 6.2 Included cases flow chart. 182

Figure 6.3 included cases demographics 183

Figure 6.4 Cases stratification according to their treatment 183

List of Tables

Table 2.1 WEB OF SCIENCE literature review protocol 61

Table 2.2 Excluded articles 62

Table 2.3 Decision support systems in Urological domain 65

Table 2.4 Diagnosis prediction application of Expert systems in Urology 71

Table 2.5 Disease stage prediction 78

Table 2.6 Treatment outcome prediction 83

Table 2.7 Recurrence and progression prediction 91

Table 2.8 Research variable prediction 96

Table 2.9 Image analysis 105

Table 2.10 Qualitative assessment of urological expert systems 109

Table 3.1 Prostate cancer risk stratification, copied from NICE guidelines 122

Table 3.2 NICE 2019 prostate cancer investigation and treatment guidelines, AS protocol. 124

Table 4.1 Decision rules change following the first panel meeting 155

Table 4.2 Kappa statistics for set 1 156

Table 4.3 Kappa statistics for set 2 156

Table 4.4 Kappa statistics for set 3 156

Table 5.1 Case stratification 164

Table 5.2 Quantitative analysis and kappa statistics estimation 169

Table 6.2 Kappa statistics for investigator 1 and CDSS 184

Table 6.3 Kappa statistics for investigator 2 and CDSS 184

Table 6.4 Kappa statistics for investigator 1 and investigator 2 184

Abbreviations

Ac	Accuracy
AI	Artificial intelligence
ANN	Artificial neural networks
AP	Acute prostatitis
Bca	Bladder cancer
BC	Backward chaining
BCF	Biochemical failure
BP	Back propagation neural network
BPD	Benign prostatic disease
BPH	Benign prostatic hyperplasia
BN	Bayesian network
BOO	Bladder outflow obstruction
CAD	Computer aided diagnosis
CBR	Case based reasoning
CNS	Cancer Nurse Specialist
CP	Chronic prostatitis
CV	Cross validation
Dom	Domain
DR	Decision rules
DRE	Digital rectal exam
ED	Erectile dysfunction
ELM	Extreme learning machine
ES	Expert systems

ESWL	Extra corporeal shockwave lithotripsy
FC	Forward chaining
FH	Family history
Fkn	Fuzzy k nearest neighbour
FLS	Fuzzy logic systems
F-ONT	Fuzzy Ontology
FNM	Fuzzy neural modelling
FRB	Fuzzy rule-based systems
FSH	Follicle stimulating hormone
GA	Genetic algorithm
Gl	Gleason score
HAT	Hormone ablation therapy
Hgon	Hypogonadism
Hk11	Human kallikrein 11
IMRT	Intensity modulated radiotherapy
INV 1	Investigator 1
INV 2	Investigator 2
INV 3	Investigator 3
IS	Information systems
ISS	Irritative symptoms
IT	Information technology
IUPA	Infundibular ureteropelvic angle
IW	Infundibular width
KA	Knowledge acquisition
KE	Knowledge engineer

KMSP	Kaplan Meir Survival Plot
kNN	k nearest network
Lap	Laparoscopy
LDA	linear discriminant analysis
LH	Luteinizing hormone
LHRH	Luteinizing hormone releasing hormone
LND	Lymph node density
LR	Linear regression
LUT	Lower urinary tract
LVQ	Learning vector quanitizer
MANN	Matlab developed ANN
MIC-1	Macrophage inhibitory cytokine-1,
MIF	Macrophage inhibitory factor,
MH	Medical history
MHRA	Medicines and Healthcare products Regulatory Agency
Mdl	Model
MLFF-BP	Multi-layer feedforward networks with backpropagation
MLP	Multi-layer perceptron
MP MRI	Multiparametric MRI scan of the Prostate
MRSI	Proton magnetic resonance spectroscopic Imaging
Nep	Nephrectomy
Nlt	Nephrolithiasis
NICE	National Institute for Health and Care Excellence
Nomo	Nomogram
NPV	Negative predictive value

Nsc	None seminoma testicular cancer
Oss	Obstructive symptoms
p2PSA	-2 proPSA
CaP	Prostate cancer
PPM	Parts per million
PPV	Positive predictive value
Pop	Pelvic organ prolapse
PSAd	PSA density
PSAv	PSA velocity
PVR	Post void residual
Qmax	Maximum flow rate
RA	Requirement analysis
RBF	Radial basis function
RBR	Rule based reasoning
RC	Radical cystectomy
RCC	Renal cell carcinoma
RCT	Randomised controlled trials
Recur	Recurrence
RELM	Regularised extreme learning machine
Res	Response
RMS	Root mean square error
RNN	Recurrent neural network
ROC	Receiver operating characteristic
RRP	Radical prostatectomy
SANN	Statistica developed ANN

Sc	Single centre
Se	Sensitivity
SPC	Stable prostate cancer
Sp	Specificity
SVM	Support vector machine
tPSA	Total PSA
Tptp	Transperineal template guided prostate biopsies
TPV	Total prostatic volume
TRUS	Trans rectal ultrasound scan
TZD	Transitional zone density
TZV	Transitional zone volume
U Dyn	Urodynamic study
U Dys	Urinary dysfunction
UTI	Urinary tract infection
V&V	Verification & Validation
VUR	Vesico-ureteric reflux
%fPSA	Percentage free/total PSA
%p2PSA	p2PSA/fPSA
x2-tests	Chi squared test

List of appendices

Appendix 1 Clinical investigation plan protocol

Appendix 2 Refined decision rules

Appendix 3 Service evaluation approval

Appendix 4 Protocol of knowledge validation study

Appendix 5 Clinical cases presented to investigators for rating in the knowledge validation study

Appendix 6 Electronic letter generated by the CDSS in the live validation study

Appendix 7 Example of cases generated for the live validation

Appendix 8 Protocol of the live validation

Chapter one - Introduction

1.1 Background

Artificial intelligence (AI) was born as a concept that machines can simulate human learning and intelligence. McCarthy, an American mathematician, was a strong believer and proposed in 1956 that 'Every aspect of human intelligence and learning can be simulated by a machine' (McCarthy et al., 2006). Indeed he was not alone as Alan Turing in Cambridge shared this strong belief and through his novel research which in the 1940s advanced the standard mathematical logic into effective computer programming (Turing, 1950) and is considered to be the father of AI.

Developing AI depends upon 2 axioms: ample and reliable memory to retain information and an ability to recognise and manipulate numbers through a predefined algorithm. These algorithms represent aspects of human cognition.

By 1975 by a collaborative group of scientists at Stanford University developed 'MYCIN' which successfully simulated the decision making of microbiologists in diagnosing and treating patients with microbial infections (Shortliffe et al., 1975). When tested, it outperformed human specialists, but because of ethical concern around using computers to make clinical decisions, it was never used in practice (Yu et al., 1979). This is one of the earliest Experts Systems (ES) to be reported in the literature and its rule based model is considered a standard model for problem solving.

In Urology ES existed in their literature as early as 1988 with a rule system called SEXPERT. It was designed to interact with couples suffering from sexual dysfunction and provide a therapeutic dialogue tailored for their individual dysfunction (Binik et al., 1988). More expert systems were developed for supporting prostate cancer (CaP) decision making. The majority of those systems applied the machine learning abilities on raw CaP data to identify patterns and multilinear relations between input variables.

This analytical ability was then applied to clinical cases and it was estimated that the use of one of those systems could prevent 68% of repeat biopsies for detecting CaP (Remzi et al., 2003). Applications of ES in CaP have also been described to: calculate radiotherapy dose (Petrovic et al., 2011), predict recurrence after surgery (Botoca et al., 2009), predict survival (Catto et al., 2009) and analysis and report of scans (Xiao et al., 2016).

However, despite the above small studies and demonstration of the potential for the widespread use of ES in healthcare as a whole (including in CaP) (Garg et al., 2005, Kawamoto et al., 2005), there have been few reports on their live testing of ES in a clinical setting (Ammenwerth et al., 2013).

NICE guidelines in 2008, 2014 and 2019 encouraged community follow up of patients after treatment for CaP, but there was resistance from primary care practitioners who felt they lacked the required expertise (Watson et al., 2011). In industry, similar problems have been successfully solved by applying ES to replicate the expert reasoning where other expertise is lacking in decision making. However, development of an ES for CaP follow up has not yet emerged as development of such systems has been shown to be complex and laborious with a poor track record of uptake in healthcare.

1.2 Expert system structure, development and validation

Expert systems can be defined as "computerised advisory programmes that attempt to imitate or substitute the reasoning process and knowledge of experts in solving specific types of problems" (Turban, 1990). They aim to simulate experts in a solving a problem which is usually domain specific (specific to a particular task). Aspects that distinguish

ES from other standard AI information systems were specified by Jackson (Jackson, 1999) and are as follows:

1. ES are programmed with knowledge acquired from the human domain expert [i.e. an expert in the field of interest, e.g. a clinician experienced in the treatment of CaP], in contrast to standard applications which are based on mathematical models or algorithms. In the example of MYCIN, the system was constructed with microbiology specialist knowledge on microorganism morphology, gram staining, infective diseases feature and sensitivity to antibiotics.
2. There is a reasoning or inferring scheme programmed in ES that processes the stored expert knowledge and the problem data presented by the user, allowing a solution to be reached [simulating the decision making of the human expert, who uses experience and knowledge to process information presented to reach a decision]. The outcome of the ES has to be similar or as close as possible to the domain expert's opinion.
3. ES apply an approximate approach in their problem solving, providing an outcome even when imperfect or incomplete data is supplied by the user. This is unlike other conventional algorithm driven programmes which require complete data sets to function, and this makes the ES similar to a human expert in applying the 'rule of thumb' in their calculation.
4. The ES should be able to explain the reasoning behind its suggested outcome and to prove that this was not by chance. For instance, MYCIN can suggest that there is a 40% chance that the patient has Salmonella endocarditis based on the presented clinical data. The system can explain this to the user by reporting the symptoms of endocarditis then their 40%

correlation with Salmonella infection. These correlations (or weights) are pre-existing knowledge programmed from the microbiologists' knowledge.

1.2.1 Expert system structure

The main objective of ES is to support the user in solving a domain specific (particular clinical area) problem, which would usually require an expert to solve. Therefore the standard design of the ES will include these three main components:

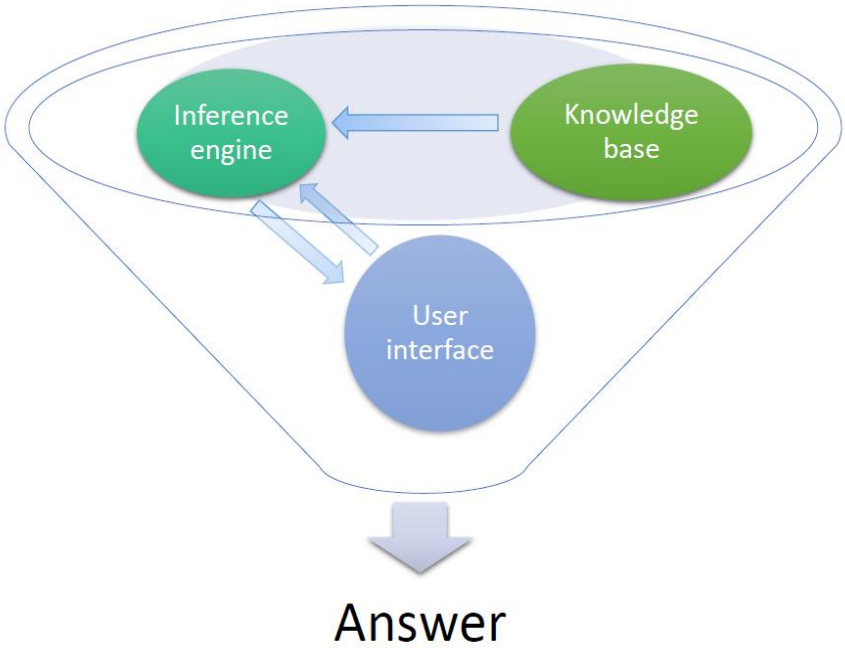


Figure 1.1 Expert system components
Three main components: knowledge base, inference engine and user interface, which interact to produce an answer in a way analogous to logical human expert processing of data.

1. *Knowledge Base:* is the core component of the system where all the domain expert knowledge is stored and manipulated for solving the problem (domain is defined as an area of interest over which the experts have control). The programmer, who is most often not an expert in the area of the ES, codifies the

knowledge into an ES framework, which requires deep understanding of the multi linear and hierarchical relations of the domain knowledge, and works by following logical operator processes (IF-THIS-THEN-THAT for example). It is here that development of ES can be considerably slowed due to the programmer's lack of expertise in the area of the ES and hence difficulty programming the system to incorporate the nuances of decision making in that domain (Pandey and Mishra, 2009, Benbasat and Dhaliwal, 1989, Achour et al., 2001).

2. *Inference engine:* is a separate programme from the knowledge base which codes for the system reasoning scheme to allow solving the domain specific problem. Unlike the Knowledge base, it is more standard in structure and not programmed by any domain specific knowledge, hence not situation dependent and similar programmes can be moved from one domain to another. Its structure (e.g. forward or backward chaining of decision making processes) is chosen by the programmer according to the system methodology.
3. *User Interface:* this allows the user to communicate with the system and to input data. Even though this component is not part of the knowledge processing programme, but it can be crucial in deciding the success of the whole system: a clunky interface will deter uptake and use of any ES.

1.2.1.1 Expert systems models:

ES can be classified by their model methodology then mapped them to their application (Liao, 2005). Liao identified eleven different ES models (Rule based, Knowledge based, Neural networks, Fuzzy logic systems, Object-oriented, Case based reasoning, System architecture development, Intelligent agent systems, Modelling ontology and Database methodology) with hundreds of applications in agriculture, environment, finance,

medicine etc. (Liao, 2005). However, there was little analysis and reflection on their development and its influence on the system model structure and development methodology. Pandey and Mishra (Pandey and Mishra, 2009) reviewed medical ES specifically, and categorised them according to their development into *knowledge dependent, data dependent and hybrid models*. Knowledge was defined as useful information that can be acquired from the domain experts, while data were identified as facts which simply exist and have no significance beyond their existence (Ackoff, 1989). Pandey and Mishra recognised that *knowledge dependent* ES use as Rule Based Reasoning (RBR) and Case Based Reasoning (CBR) and have a separate knowledge base for storing the domain expert knowledge. This component is not required for the data dependent (*intelligent computer models*) systems such as artificial neural networks (ANN) and Fuzzy Logic (FL) which rely on mathematical models for estimating relations between relevant clinical data then link them to an outcome (i.e. do not use expert knowledge, only statistical association, to suggest an answer).

1.2.1.2 Functional design of some common models:

1.2.1.2.1 Rule based reasoning:

RBR is considered a standard structure for problem solving in AI (Luger, 1998), an example of this is the early MYCIN system, where its knowledge was represented in the

form of conditioned phrases (IF-THEN) known as 'symbolic rules'. A rule example from MYCIN:

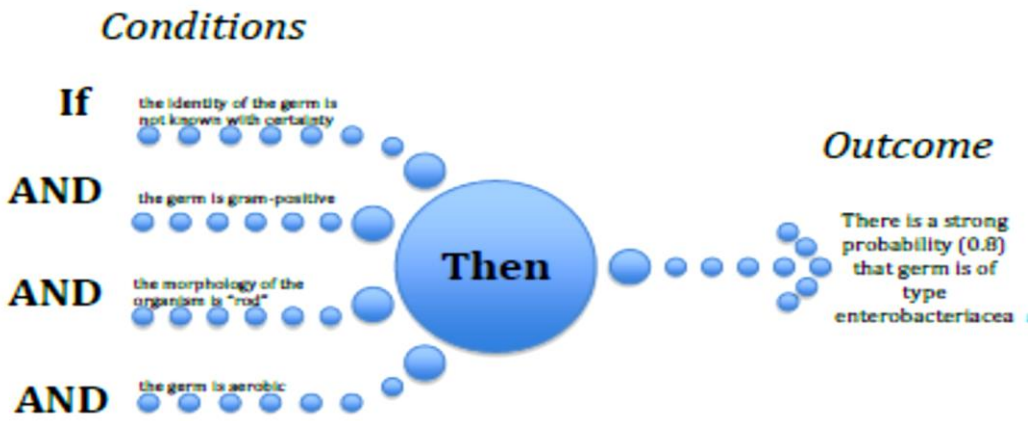


Figure 1.2 Example of a rule from MYCIN

A question is generated by the user through the interface, which is then translated into rule conditions (IF, AND, AND) (Figure 1.2) and stored onto the knowledge base. The inference engine will subsequently map these conditions to the stored rules and navigate through the knowledge base in order to generate an answer (…THEN).

There are two standard methodologies that the inference engine can apply. The first one is *forward chaining* where the system starts the reasoning from the input data (clinical facts) and map them to the stored rules on the knowledge base until a rule is triggered. An example in the diagram below were the system has facts (1, 2&3) and rules to choose from.

Rules:

Rule 1 (*IF* Fact **a** + Fact **b** + Fact **c** *THEN* Ouctome **X**)

Rule 2 (*IF* Fact **c** + Fact **d** *THEN* Ouctome **Y**)

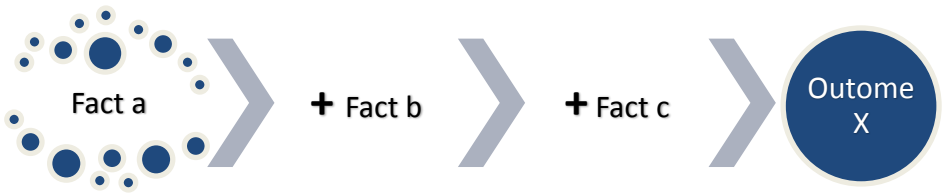
Cases:

Case 1: problem with fact **a** + fact **b** + fact **c**?

Case 2: problem fact **c** + fact **d**?

Forward chaining from conditions (facts) to outcome (output)

Equation 1: Forward chain reasoning through case 1



Equation 2 Forward chain reasoning through case 2

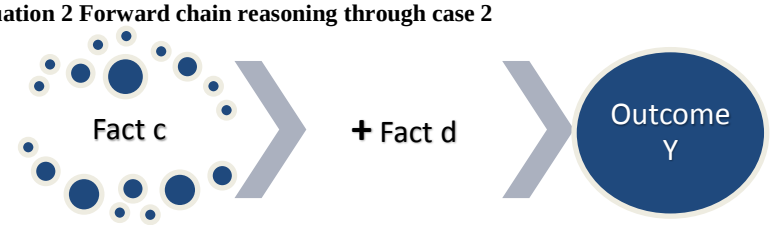


Figure 1.3 Example of Forward chaining

In case 1: the system will start with the facts as in case 1 adding them together then search for a rule with similar fact in its conditions. This was matched to rule 1 therefore outcome X was triggered. Similarly, case 2 had outcome Y.

Rules:

Rule 1 (*IF* Fact **a** + Fact **b** + Fact **c** *THEREFORE* Ouctome **X**)

Rule 2 (*IF* Fact **c** + Fact **d** *THEREFORE* Ouctome **Y**)

Cases:

Case 1: fact **a** + fact **b** + fact **c**?

Case 2: fact **c** + fact **d**?

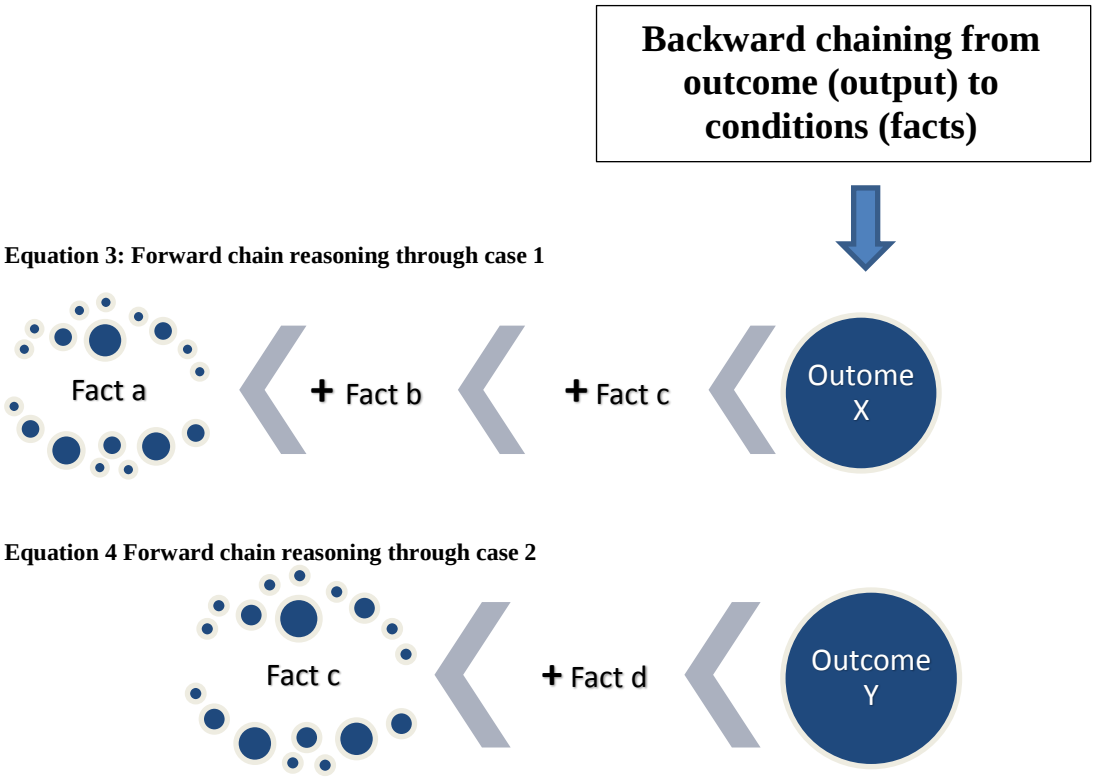


Figure 1.4 Backward chaining
The reasoning starts from the outcome then the machine will search in the case facts until its rule (s) is satisfied. Otherwise a different outcome will be selected then further search will take place until a rule is satisfied.

In forward changing, the system will start with the facts as in case 1 adding them together then search for a rule with similar fact in its conditions. This was matched to rule 1 therefore outcome X was triggered. Similarly, case 2 had outcome Y. The opposite is applied in backward chaining where the system hypotheses that one outcome is correct then search in the case for facts matching its rule (s) until a rule is found or this hypothesis is disputed and a different outcome will be chosen to repeat the same

process again. This will continue until a rule is satisfied and the outcome is then confirmed.

Even though the forward chaining might seem more logical method for reasoning, the backward chaining is more common (Pandey and Mishra, 2009). In fact, backward chaining is very common in human specialist clinical decision making. An example is patients referred with high PSA. The specialist hypothesis that this patient has prostate cancer and then spends the rest of the consultation eliciting relevant clinical facts such as age, family history, persistently elevated PSA and abnormal feeling prostate on digital rectal exam (DRE) to test his hypothesis. If those clinical data are present, then a rule in his clinical knowledge base is satisfied and a diagnosis of suspected cancer is made. Alternatively, if the PSA is declining and the patient has had a confirmed infection on urine culture then the specialist may shift his reasoning to an alternative diagnosis of prostatitis and dismiss prostate cancer as the decision rules to suspect prostate cancer were not satisfied. In this scenario, the clinician reasoned backwards by starting from suspecting the diagnosis (expert system outcome) which is similar to the machine backward chaining. In the same scenario the primary care physician has probably applied forward chaining as they started from just one clinical fact which may have been just urinary frequency. The physician then have to reason forward by eliciting more relevant clinical data which have led to the performance of the PSA and referral.

1.2.1.2.2 Artificial neural networks (ANN)

ANN are computational methodologies capable of performing multifactorial analyses inspired by the biological structure of the nervous system (Anagnostou et al., 2003). As in the nervous system, their network has layers of interacting nodes, connected by axons and synapses (Figure 1.5). The nodes represent the domain input and output variables

which the system has identified in the machine learning phase. The network synapses (between input and output nodes) represent weighted mathematical connections which the systems have also developed and refined in the machine learning phase. This complex relation can be further demonstrated by dissecting the network to examine its functional elements which is similar to examining the individual nerve cell.

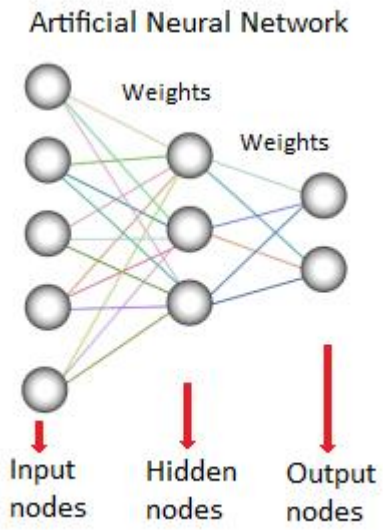


Figure 1.5 Artificial neural network structure
Hidden layer have adjusted weights acquired from the machine training to solve domain problems

The first neural network was developed in 1959 by Rosenblatt. This had a simple structure of linearly arranged nodes (or variables) in two distinct layers of input and output nodes (Rosenblatt, 1958) .The system reasoning starts with the node value being multiplied by the connection weight and when result is above the threshold, the output node will be triggered and a conclusion can be drawn (Figure 1.6).

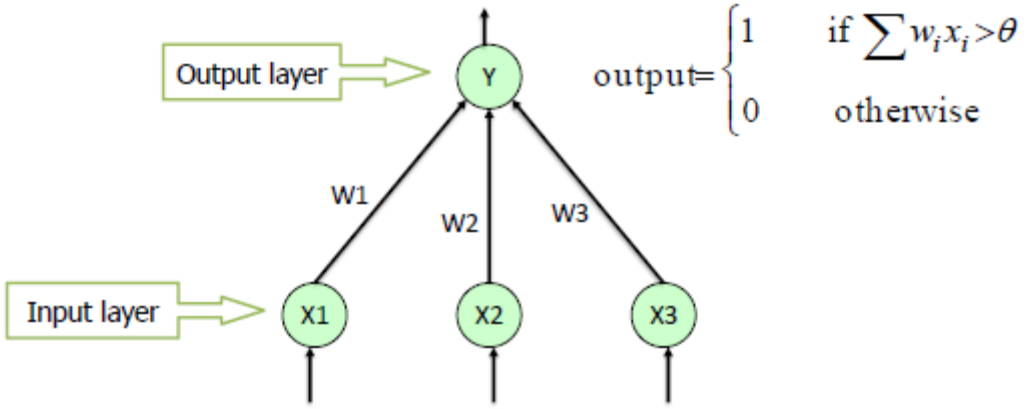


Figure 1.6 Single perceptron structure
The sum of the inputs (X_i) are multiplied by the sum of the weights (W_i). If the sum was above the threshold then outcome one is triggered otherwise the outcome will be 0.

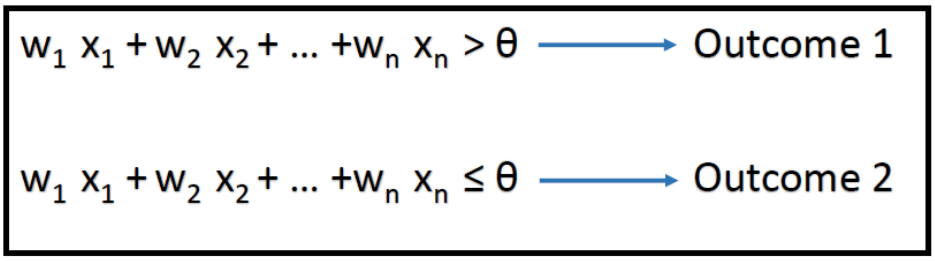


Figure 1.7 Perceptron input nodes weighting to trigger outcome
As mentioned above in figure 1.6, the sum of the inputs x weights assigned by the machine for each nodal connection is checked against the threshold to trigger an outcome.

In contrast to the knowledge base, ANN relies on mathematical formula rather than symbolic logic as in the knowledge based systems. This mathematical formula has a standard structure as in Figure 1.6 Single perceptron structure) with input nods and output. Weights are calculated and adjusted in the learning phase from real data. A simplified example of the concept of machine learning is mentioned below and the figures are all approximation for illustration only.

Example of machine learning from clinical data of patients with known prostate cancer confirmed by histology with aim to develop a predictor of cancer based on clinical and lab data as family history, age and PSA level. The defined output will be cancer or no cancer.

Inputs

1. Family history (FH)

   no FH = 0
   1 member = 0.1
   2 members = 0.2
   3 ……..

2. Age

   70 = 0.7
   75 = 0.75
   80…..

3. PSA

   15 = 0.15
   20 = 0.20
   2……

Outputs

1. Cancer if sum is > 1 (threshold = 1)
2. No if sum < 1 (threshold = 1)

Equation

Weight x input 1 (FH) + Weight x input 2 (Age) + Weight x input 3 (PSA) = Y

If Y > 1 then cancer

If Y < 1 then no cancer

Machine training with cases of known outcomes (output)

In the following cases the machine will use 1 as weight

- Case 1. Family history of CaP (1 family member = 0.1), age (75 = 0.75) and PSA (15 = 0.15). Machine will leave the weight at 1 for all then the equations:
  - *(FH x 1) + (Age x 1) + (PSA x 1) = 0.1 + 0.75 + 0.15* = 1 = threshold
  - Therefore output is **CANCER** then weights are acceptable

- Case 2. Family history of CaP (1 family member = 0.1), age (65 = 0.65) and PSA (15 = 0.15). Machine will leave the weight at 1 for all then the equations:
    - *(FH x 1) + (Age x 1) + (PSA x 1) = 0.1 + 0.65 + 0.15 = 0.9 < than 1*
    - Therefore output is **NOT CANCER** which is the wrong answer as this is a known case of CaP.

The machine will then reiterate the weights to improve predictability, in this example it will assign a higher weight to FM to be 2 instead of 1.

- Case 1: family history of CaP (1 family member = 0.1), age (75 = 0.75) and PSA (15 = 0.15).
    - *(FH x 2) + (Age x 1) + (PSA x 1) = 0.2 + 0.75 + 0.15 = 1.1 > threshold*
    - Therefore output is **CANCER** then weights are acceptable
- Case 2. Family history of CaP (1 family member = 0.1), age (65 = 0.65) and PSA (15 = 0.15).
    - *(FH x 2) + (Age x 1) + (PSA x 1) = 0.1 + 0.65 + 0.15 = 1 = threshold*
    - Therefore output is **CANCER** then weights are acceptable.

Figure 1.8: An example of a single layer perceptron machine learning on cases with known prostate cancer
The clinical variables of patients' age, family history and PSA levels were used as inputs for the machine. A set of weights were then applied assuming that all variables have the same weight. Their accuracy was then improved when the family history had more weight than the other variables.

The mathematical model is refined until a predefined level of accuracy is reached (sensitivity and specificity) then the machine will be tested with new set of data to estimate its true accuracy.

The above example described the single layer perceptron which is only capable of performing linear reasoning as they rely on linear mathematical equation. This is not effective for multilinear relations which lead to the development of multi-layered perceptron (MLP) by Werbos (Werbos, 1974). This MLP added an extra layer of nodes

between the inputs and outputs which allowed the machine to represent the multilinear relations between the variables in multiple interlinked nodes. If this is applied to the example in Figure 1.8, the machine can learn more complex patterns and improve its prediction power. However, it made the calculations very ambiguous and impossible to understand or predict. This is called the 'black box phenomenon' and it is this lack of trust in the neural networks that has challenged its acceptance in the medical field (Pandey and Mishra, 2009, Catto et al., 2009, Papageorgiou, 2012).

1.2.1.2.3 Fuzzy logic systems (FLS)

FLS are systems applying the fuzzy sets concepts to model uncertainty in medical domains (Papageorgiou, 2012, Olufunke et al., 2013). Arising from Set Theory, the fuzzy set is 'a class of objects with a continuum of grades of membership' (Zadeh, 1965). They aim to reduce uncertainty to a single figure in order to simplify the reasoning and reduce the data load input to the machine. This concept is not uncommon in clinical practice, for example, the early warning score in common use on wards, where the multiparameters of the vital signs are all reduced to a single figure to indicate the patient status and triggers a predefined action. Fuzzy logic set membership function assigns to each input a degree of certainty ranging from zero to one for each input as in the example below reducing the body core temperature into one of three categories (Figure 1.9).

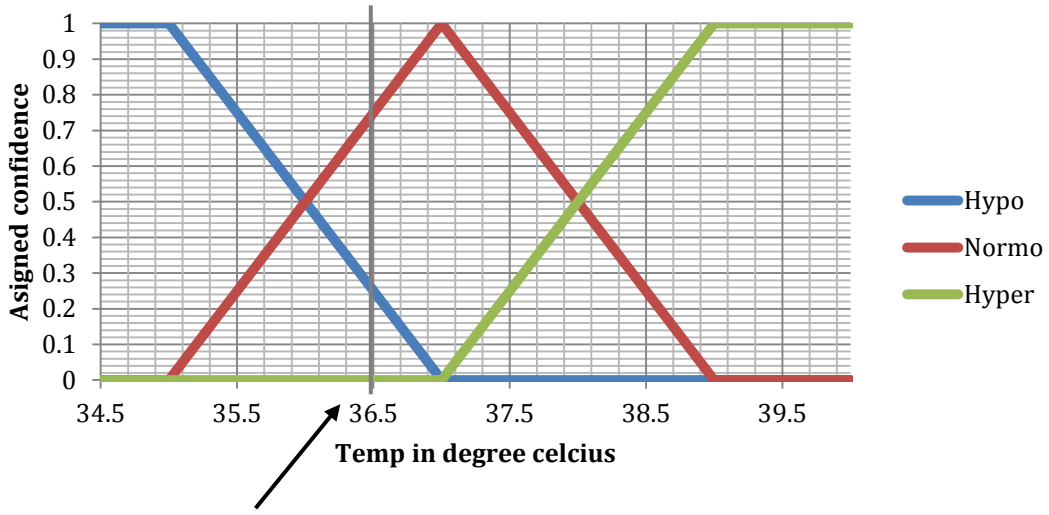


Figure 1.9 Fuzzy Logic modelling

The graph has a horizontal axis with the core temperature score in degree Celsius and vertical axis with a weight ranging between 0-1. The three coloured lines intersect with the vertical line at 36.5 °C at 3 points with a weight value on the vertical line of 0, 0.25 and 0.75. The highest weight is 0.75 on the normothermia line, therefore the temperature of 36.5 °C is normothermia with 0.75confidence. Moving the vertical line to another temperature will produce different intersections and different weights: at 38.5 °C for example, the green line has the highest weighting and hyperthermia is the output.

The function of this membership set is to represent the numerical uncertainty of the core body temperature in Degrees Celsius to one of three classes (normothermia, hypo – hyperthermia) with a degree of confidence. For instance, the grey line representing temperature of 36.5 °C has met the three graph line at three different points: hypothermia at 0.25, hyperthermia at 0 & normothermia at 0.75. This will be transcribed by the system that a core temperature of 36.5 °C is normothermia with 0.75 confidence as it is higher the hypothermia which is only 0.25 confidence. This lines are pre-set by the expert.

While this might seem too simple it can significantly reduce the need for a large number of rules to represent each numerical figure of the as in the example above where just one membership function can classify any number between 34 to 42 degree Celsius to a figure between 0 and 1.

1.2.1.2.3.1 Fuzzy Rule based systems (FRB)

Fuzzy logic can be combined with a rules based system. The application of this fuzzy logic to the RBR can reduce the number of rules in the knowledge base and simplify the system. Olufunke et al. reported that the introduction of fuzzy logic reduced the rule base by 63% while maintaining the accuracy at 90% for estimating the coronary heart disease risk in non-smoking men (Olufunke et al., 2013).

Example of Fuzzy Rule Base rule:

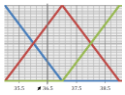
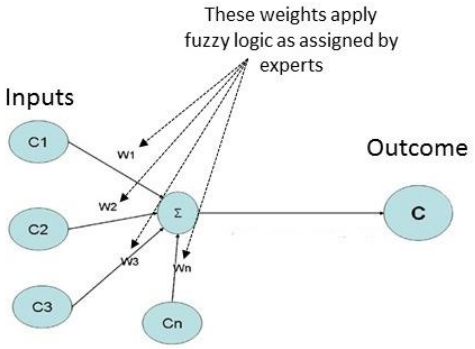
IF  0.75 Normothermia *THEN* sepsis unlikely

Figure 1.10 Fuzzy rule base
In this example, the input value of the core temperature of 36.5 °C has been replaced by the membership function which applied the fuzzy logic and produced a crisp value of 0.75 weight for normothermia. The later was then used as an input for the IF – THEN rule to trigger an outcome: IF 75% chance of the patient being normothermic at 36.5C THEN sepsis unlikely.

1.2.1.2.3.2 Fuzzy neural modelling (FNM)

FNMs have a similar structure to ANN with nodes representing variables for inputs and outputs (Pandey and Mishra, 2009, Catto et al., 2009, Papageorgiou, 2012). As explained in 1.2.1.2.2, ANN has an added layer of nodes between the inputs and outputs to allow the non-linear variable analysis. In FNM, this has been replaced by weighted connections applying the Fuzzy Logic (Papageorgiou, 2012, Catto et al., 2009). Thus on activation of the input nodes, the weights are calculated by the fuzzy function model of their individual connections, are added together to trigger the output node if above the threshold value (Papageorgiou, 2012) (Figure 1.11).



$$w_1 x_1 + w_2 x_2 + \ldots + w_n x_n > \theta \longrightarrow \text{Outcome 1}$$

$$w_1 x_1 + w_2 x_2 + \ldots + w_n x_n \leq \theta \longrightarrow \text{Outcome 2}$$

Figure 1.11 Fuzzy neural modelling
Unlike neural network, the weights of the connections are run by fuzzy logic to quantify the uncertainty. Those weights are assigned by the domain experts in developing the FL membership function and tested until system is accurate.

Example of Fuzzy neural modelling on clinical data of patients with prostate cancer confirmed by histology with aim to develop a predictor of cancer based on clinical and lab data as family history, age and PSA level. The defined output will be cancer or no cancer.

Equation of FNM

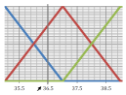
Input 1 (FH) x weight 1 +input 2 (Age) x weight 2 + input 3 (PSA) x weight 3 = Y

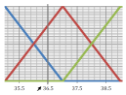
If Y > 1 then cancer

If Y < 1 then no cancer

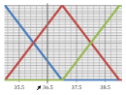
Inputs

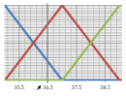
1. Family history (FH)

1 member X  = 0.1

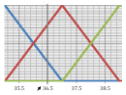
2 members X  = 0.2

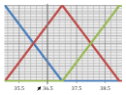
2. Age

70 X  = 0.7

75 X  = 0.75

3. PSA

15 X  = 0.15

20 X  = 0.20

Outputs

1. Cancer if sum is > 1 (threshold = 1)
2. No cancer if sum < 1 (threshold = 1)

Cases:

- Case 1: family history of CaP (1 family member = 0.1), age (75 = 0.75) and PSA (15 = 0.15):
  - *(FH x w1) + (Age x w2) + (PSA x w3) = 0.1 + 0.75 + 0.15* = 1 = threshold
  - Therefor output is **CANCER**

Figure 1.12 An example of fuzzy neural modelling calculations
The clinical variables of patients' age, family history and PSA levels were used as inputs. Those inputs are then multiplied by the weights generated by the membership function which are predefined by the experts. Those member ship function apply the fuzzy logic and are more transparent than the neural network hidden layer of nodes.

Unlike the neural network hidden layer, the weights in FNM are more transparent and easier to understand and predict. As assigning a higher weight to family history of CaP in the cancer prediction model would seem more appealing to the Urologist than a string of numbers in a mathematical equation. On the other hand, this human dependant process are dependent on overcoming the challenges of knowledge elicitation from the domain experts (Pandey and Mishra, 2009).

1.2.2 **Expert systems development life cycle:**

ES are intelligent information systems (IS) capable of simulating the domain expert in solving a specific problem. Their development is formed of nine integrated stages: initiation, system concept development, planning, requirement analysis, design, development, testing, implementation, maintenance, and disposition (Ashrafi et al., 1995). This can be further simplified into five main stages: analysis, design, implementation, testing and maintenance (Figure 1.13) (Ashrafi et al., 1995). However, ES distinctive features (mentioned in 1.2.1) have led to the evolution of most of the IS development stages in order to adapt to their unique system requirements.

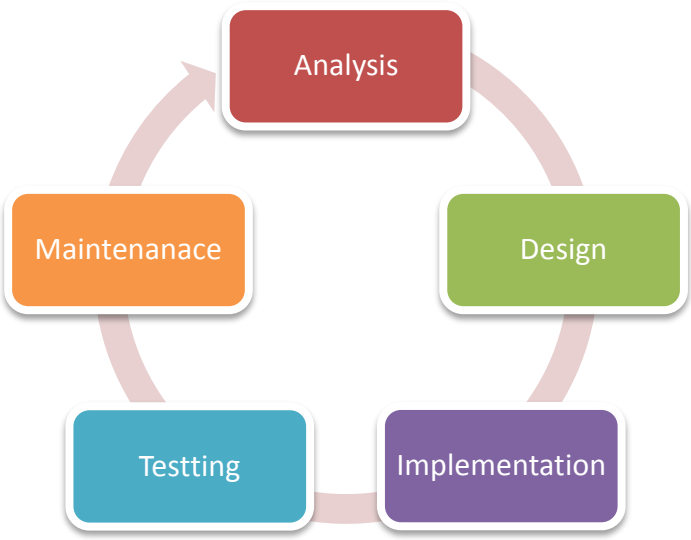


Figure 1.13 Information system development cycle
These are the main stages in the information system development. Those have been modified to meet the requirements for expert systems development specially the analysis and testing phase.

The first and probably the most essential step for developing a useful IS, is the "Requirement analysis" (RA) where an analyst gathers information about the end user need, identifies the system objectives and plan the system design (Byrd et al., 1992). In ES, it is clear that an analyst will face a significant challenge in eliciting complex

domain knowledge from the experts. Therefore, the standard techniques for RA are not sufficient to perform this task and it has been replaced by knowledge acquisition (KA) (Byrd et al., 1992). The KA will be discussed below in 1.2.2.1.1.1. It is considered the bottleneck for ES development and ES development is likely to fail at this stage than any other (Pandey and Mishra, 2009, Okeefe et al., 1987, Benbasat and Dhaliwal, 1989).

Furthermore, the knowledge base structure, reasoning pattern and ES modelling requires a more complex design development and testing than standard IS. This is more prominent in medical ES development where the whole system has to go through rigorous testing in order to meet the standards for registration with the regulatory authorities such as the Medicines and Healthcare products Regulatory Agency (MHRA) in the UK (Ammenwerth et al., 2013).

1.2.2.1 Knowledge based development:

The knowledge base is the key component of the ES, storing all the essential domain expert knowledge for solving the user's problem by inference. It is predominantly built with human knowledge, which is more complex to codify into programming logic than standard information based algorithms. Thus, the requirement analysis (RA) in the IS life cycle has been substituted by the knowledge acquisition (KA) to facilitate eliciting knowledge from the domain experts and developing the knowledge base.

1.2.2.1.1 Knowledge acquisition (KA):

This is a process where knowledge is acquired from the human domain experts and scientific evidence to construct the knowledge base (Grosan and Abraham, 2011). Its main objectives are *eliciting*, *acquiring* and *representing* knowledge of descriptions, relationships and procedures in a specific domain (Benbasat and Dhaliwal, 1989). The knowledge engineer plays an important role by catalysing the knowledge transfer from

the domain expert to the system developer in order to construct the knowledge base. It can be challenging (Pandey and Mishra, 2009, Benbasat and Dhaliwal, 1989, Achour et al., 2001, Okeefe et al., 1987) due to the following reasons:

1. The knowledge is usually presented to the knowledge engineer in its natural technical language (jargon) who may well not be familiar with the domain and hence the jargon. This knowledge has to be digested by the knowledge engineer before being encoded by the system developer (Kidd, 1987).
2. The domain expert may not be able to express the knowledge effectively, specially the tacit (implicit or subconscious) knowledge which has been gained through years of experience and processed in the experts' subconscious (Gebus and Leiviska, 2009).
3. The possible lack of the required level of expertise and unwillingness to share their knowledge (Gebus and Leiviska, 2009).
4. Further complexity of the medical knowledge, which combines cause-effect structure, with human experience, empirical facts and scientific data (Boegl et al., 2004).
5. The knowledge engineer has to understand the domain problem in order to express the knowledge into their correct category: *factual knowledge* (static type describing characters of facts), *inferring knowledge* (describes the reasoning scheme required to analyse these facts to a meaningful and correct outcome) and *strategic knowledge* (it defines the hierarchical relation of facts) Figure 1.14 (Schreiber et al., 1994)

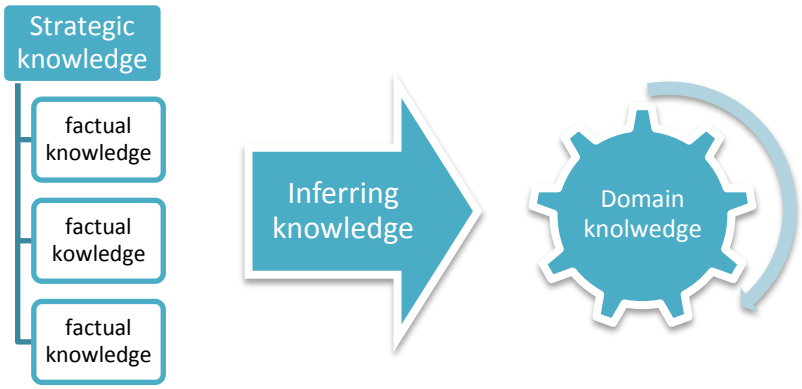


Figure 1.14 Types of knowledge
The knowledge can be stratified into three categories: strategic, factual and inferring

An example for classifying medical knowledge on localised prostate cancer:

1. Localised prostate cancer is confirmed by MRI
2. Localised prostate cancer is treated with radical surgery
3. Surgery is not ideal for elderly patients above 80.

Each of these statements represents a static fact thus can be classed as *factual knowledge*. The first two can be linked by the IF-THEN clause which consolidated the relation between the facts and this can be classed as *inferring knowledge*.

IF a patient has localised prostate cancer on MRI

THEN radical surgery is an option.

Figure 1.15 Rule based reasoning
IF – THEN are linking two facts, therefor this clause is "inferring knowledge" while the static facts are "factual knowledge".

However, this rule combining the first two facts together contradicts the third statement. Thus, a third type of knowledge is required *hierarchical knowledge* to define their ranking and in this example the fact that the patient is elderly will preclude the indication for surgery.

Some authors categorises knowledge as *declarative* and *procedural* knowledge (Beligiannis et al., 2006, Macdonald and Witten, 1989, Fujihara et al., 1997) with the

declarative being the static or factual type (similar to factual knowledge) and procedural the more dynamic (combining both inferential and strategic) that enables experts to perform the task or solve problems.

1.2.2.1.1.1 The knowledge acquisition objectives

These can be summarised in three main steps:

1. *Knowledge elicitation*: is the process where the knowledge engineer extracts the knowledge required by the domain expert to solve a specific problem or perform a task. This can be achieved by direct interaction through structured or unstructured interviews, observational task analysis or be automated using another expert system, which will be discussed below in the KA models.
2. *Knowledge documentation*: is the knowledge structuring, analysis and recording from the domain expert or scientific facts before being represented in the knowledge base (Turban, 1990). In this step, the knowledge engineer defines the relation between the knowledge fragments elicited from the domain expert in an explicit framework. This provides a structure of the elicited knowledge, which can be reviewed and corrected by the domain experts and allow the knowledge engineer to decide the appropriate model and strategy for representing the knowledge (Enting et al., 1999).
3. *Knowledge representation*: in this step, the knowledge engineer forms explicit concepts on the documented domain knowledge and codifies them in an appropriate representation scheme to achieve the target performance accuracy (Gaines and Shaw, 1993). The formalisation of these concepts requires deep understanding of the domain knowledge. To achieve this

understanding, knowledge engineer use techniques of learning (Fujihara et al., 1997, Macdonald and Witten, 1989):

a. Inductive learning to identify the patterns of reasoning used by the expert to perform a task or solve a problem (inferential and strategic knowledge of the domain)
b. Deductive learning techniques for establishing the factual domain knowledge.

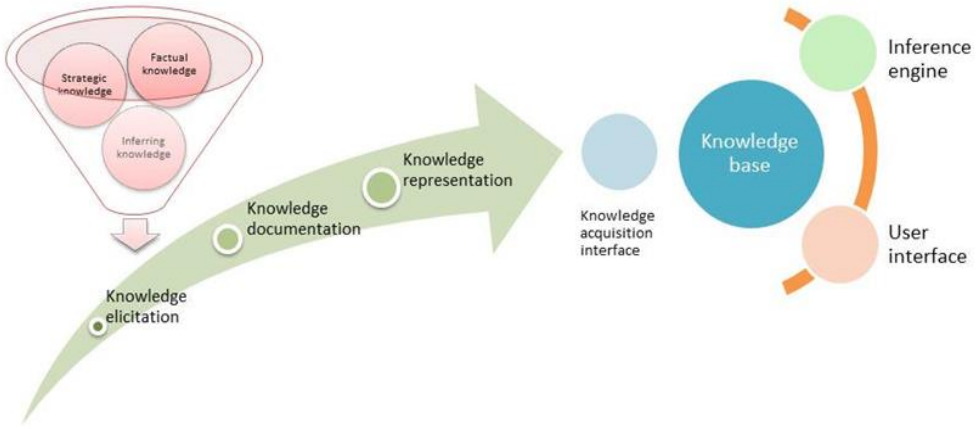


Figure 1.16 Knowledge acquisition steps and objectives
The Knowledge engineer elicits the knowledge from domain experts and classes them into factual, strategic and inferential and then uses them to structure the knowledge base through the knowledge acquisition interface.

1.2.2.1.1.2 The knowledge acquisitions models:

The objectives of the KA are to elicit, document and represent the knowledge in a functioning framework allowing the system to solve domain specific problem or perform certain tasks. The published models for KA demonstrated a diversity of their methodologies and can be categorised as follow:

1. *Direct methods* (Olson and Rueter, 1987): in this methodology, the domain expert reports their own experience and knowledge directly to the KE models, examples:
    a. Structured and unstructured interviews
    b. Questionnaires
    c. Observation of the task performance: the KE observes and records the reasoning followed by the domain expert while performing the task, which will be reviewed later to reflect on the underlying knowledge.
    d. Protocol analysis: similar to the observation but with a 'think aloud' approach from the domain expert explaining their goal at each step of the task performance.
    e. Inferential flow analysis: the KE identifies a group of key variables in solving the problem and the domain expert will define the relation between these variables.
2. *Indirect methods* (Olson and Rueter, 1987, Benbasat and Dhaliwal, 1989): the knowledge engineer infers from the tasks performed or cases solved by the domain expert to reveal the underlying reasoning scheme and tacit (subconscious) knowledge which the domain expert could not articulate in the direct interaction. However, their use is limited and are often used to direct other methods such as structured interviews (Yang, 1995).
3. *Prototyping:* in this approach, an early prototype of the system is developed to be used by the domain expert as a tool for KA and error correction at early stages (Yang, 1995). However, it can be criticised for having to decide the reasoning strategy and knowledge representation

based on immature knowledge and concepts gained from provisional knowledge acquisition. For instance, if the reasoning framework is based on rules (IF-THEN) then all acquired knowledge have to be interpreted as rules or otherwise ignored which could lead to shallow knowledge.

4. *Automatic tools:* This model applies the concept of machine learning where the domain expert interacts with KA software rather than the knowledge engineer. These tools are AI programmes, which have the ability to learn the reasoning pattern of the human expert and express this explicit knowledge in the form of symbolic (IF-THEN) or none symbolic phrases. This approach can decrease the workload of the knowledge engineer and avoid the possible communication difficulties with the domain expert; however, they have limited knowledge eliciting power and may only result in an incomplete knowledge base (Yang, 1995).

1.2.2.2 Expert system testing

Quality assurance process of information systems development, known as the Verification & Validation (V&V) process was adapted to test and validate ES (Okeefe and Oleary, 1993). Despite being an essential step of the system development, there is a wide lack of formality in the methodologies used (Okeefe et al., 1987) with this gap being more evident in the medical expert systems (Ammenwerth et al., 2013). This has limited the ability of ES to meet the high standards for testing of medical devices and lead to the lack of their uptake in health care (Ammenwerth et al., 2013).

Despite the lack of consistency, a review by O'Keefe et al. described five common stages for ES V&V: (Okeefe and Oleary, 1993)

1. Credibility (how the system is perceived by the user).

2. Evaluation (system usability).
3. Assessment.
4. Validation (building the right system).
5. Verification (building the system right).

1.2.2.2.1 Expert system validation:

Validation is 'building the right system' i.e. a system that solves the problems correctly. It is a crucial stage of the quality assurance process in the system life cycle where the system output and calculations are checked against the predefined standards in order to be deemed valid (Okeefe and Oleary, 1993). Those standards are usually described by the domain experts for solving the domain specific problem. Therefore, the main question in this step "Is the system as accurate as the expert?" (Okeefe and Oleary, 1993, Okeefe et al., 1987, Lee and Okeefe, 1994)

An example of this framework was suggested by Knauf et al. based on the Turing testing to evaluate knowledge based systems (Knauf et al., 2002):

1. *Test case generation:* a sufficient number of test data are generated to trigger all potential inputs and subsequent outputs in the format of test cases (test data + output = test case). The test case representation needs to be balanced between covering all possible cases and generating adequate statistical power, while maintaining a realistic prospect of the analysis. For instance, a system having 1000 rules may generate more than 10,000 possibilities and it is not practical to generate and test all of these cases.
2. *Test case experimentation:* it is an evaluation of the system response to the test data against human expert(s) while minimizing the bias as much as possible.

3. *Evaluation:* the results of the experimentation phase are evaluated to identify the knowledge or system errors.
4. *Validity assessment:* it is the formal validity report of the whole system based on the analysis of the evaluation report.
5. *System refinement:* the errors can be corrected in this section and the process is repeated from the relevant step.

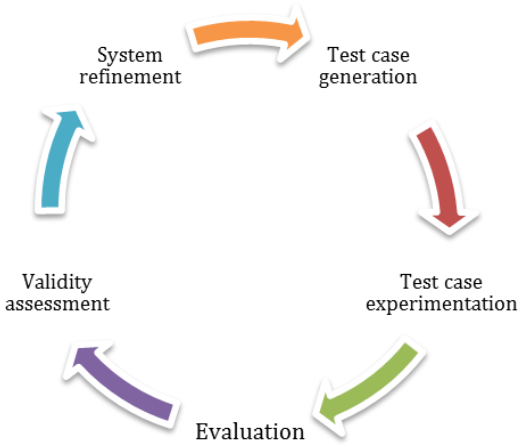


Figure 1.17 Knauf et al. expert system five stage system validation cycle
The validation relies on generating test cases triggering all knowledge fragments. Errors are then identified before the system is refined and the testing cycle is repeated until reaching the target validity point.

1.2.2.2.2 Expert system verification:

Expert system verification is 'Building the system right' (Okeefe and Oleary, 1993) and it can be defined as " the demonstration of consistency, completeness and correctness of the software at each stage of the development life cycle" (Adrion et al., 1982). Unlike validation, it is presumed that the pre-defined knowledge (acquired from the experts) and design specifications are correct and the main aim is to demonstrate consistency and correctness. Indeed, some models of verification are even independent from the knowledge, which will be discussed below.

As in other stages of the development cycle, the verification had to be modified to adapt the knowledge based systems and in a review of the literature by Okeefe et al. , these following methodologies have been reported (Okeefe and Oleary, 1993):

1. *Domain dependent verification*: called domain dependent since it relies on Meta knowledge (knowledge about the domain knowledge) of the domain which has been acquired at early stages of the KA. This meta knowledge describes and infers on the concepts and knowledge used by the system (Davis and Buchanan, 1977). This meta knowledge can be set out as concepts to be used in for verifying the systems after the validation For instance, if the meta knowledge include a concept that an elderly patient should not be offered curative surgery (as in example on page 44), then any rule contradicting with this meta knowledge concept should be checked for correctness.
2. *Domain independent verification:* this model does not presume any previous knowledge on the system domain, instead it applies the concept of anomaly detection in the knowledge representation structure.
3. *Errors in the inference engine:* checks are usually done before verifying knowledge base, as they are usually more standard in structure.
4. *Automated tools:* The whole verification process can be run by software that can verify the system by applying the domain dependent concept (Chang et al., 1990) or domain independent methodology (Preece et al., 1992).

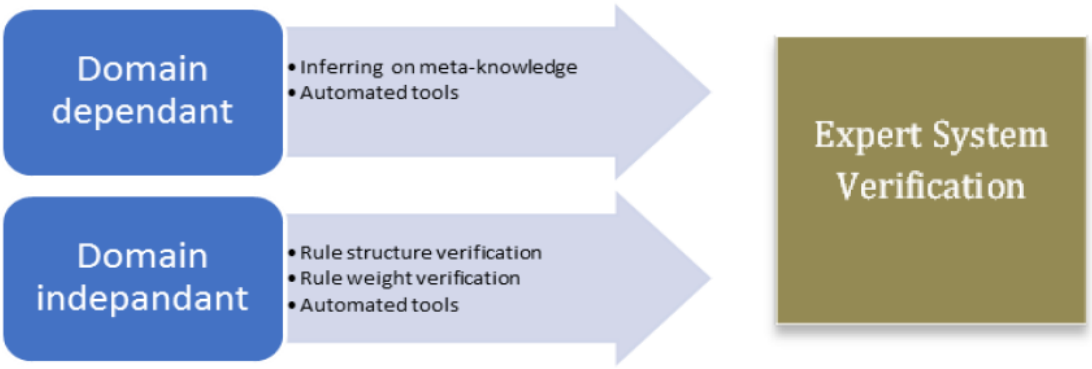


Figure 1.18 Expert system verification models
Models can be subdivided according to their aim into domain dependant and independent. The former relies on domain specific knowledge while the later focus on correcting anomalies

1.2.2.2.3 Expert system evaluation:

In computer science system, evaluation means performance time optimisation. However, in medicine it usually involves a comparison between the system performance and a panel of human experts (Lundsgaarde, 1987). The system evaluation of ES has different objectives from the above with more focus on testing the system usability rather than the error correction and output accuracy, which are the domain of the V&V part of the cycle. It is recognised to be a crucial step of any system development that can predict its failure or success (Okeefe and Oleary, 1993). For instance, a system that is too slow or with complex user interface is more likely to face user rejection despite producing the correct outcomes (Okeefe and Oleary, 1993).

Evaluation can be stratified into (Benbasat and Dhaliwal, 1989) (Figure 1.19):

a. *Evaluation of quality:* a domain expert evaluates the quality of the system output. For example, if the device output is in the form of written advice to patients, then the language must be clear, concise, of adequate quality and in their natural language.

b. *Evaluation of utility*: end user evaluation of the perceived benefits of using this system. For instance, a system designed to support primary care practitioners to identify high risk patient for developing coronary disease, should demonstrate the benefits as decreasing the consultations time and identifying the high-risk patients.

c. *Evaluation of usability*: establishes that the system is easy to use by the average target user.

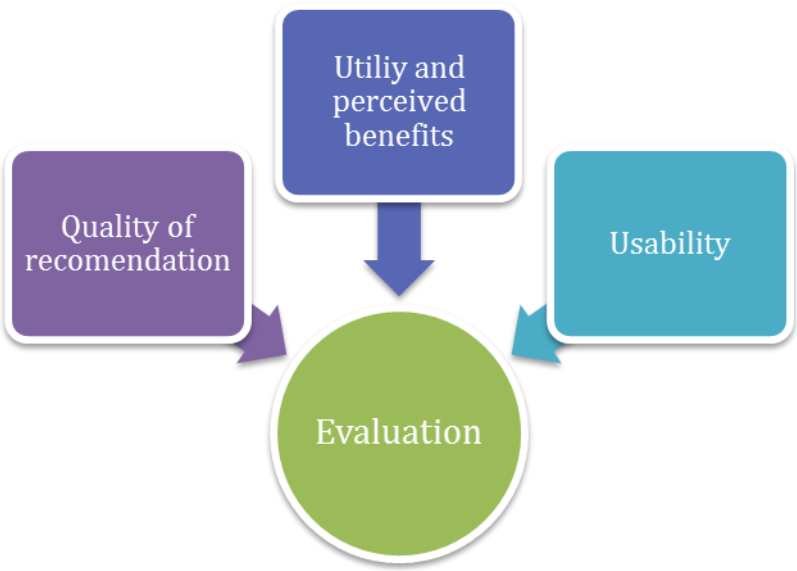


Figure 1.19 Expert system evaluation
Three main domains, quality of the output, utility (value of using the system) and the system being easy to use by the standard end user

1.3 Discussion

Application of ES can be highly beneficial to patients' care as demonstrated by Remzi et al. (2003), reporting that ANN that could reduce up to 68% of unnecessary repeated TRUS biopsies to diagnose CaP and more that are discussed in chapter two. Feasibility of their application in practice is limited by two main factors. First is the ethical and professional approval of neural network in the decision making of cancer care then their ambiguous reasoning (black box phenomenon).

The domain of prostate cancer diagnosis is itself also limiting due to its complexity and being predominantly driven by multi-disciplinary team approach. Other domains as stable prostate cancer follow up can be more appropriate for systems replicating specialist due to the need for experts. ANN may still struggle for approval due to the above mentioned factors but a knowledge based system may have some advantages due to explicit reasoning and knowledge schemes. This will come at a large cost of the long and exhaustive knowledge acquisition from experts and scientific evidence followed by rigorous testing for registration. If successful it can be used as replication of expert follow up or adjunct to other non-expert users for community follow up.

1.4 Summary

Expert systems have a very wide range of models and applications which can allow them to perform complex function for simulating human experts. They lack a standard approach for their development and testing which has limited their implementation in health care. Knowledge base systems have an exhaustive and long development but can be less challenging in their validation and registration with the regulatory bodies due to their explicit structure and reasoning. In contrast, ANN can learn highly complex patterns directly from data, but their ambiguous calculations despite their accuracy, have led to their rejection in health care.

1.5 Aims and objectives of this thesis

- To evaluate the prevalence and development of Expert Systems in the scientific literature and evaluate their strength and shortcomings.

- To formulate a development and validation life cycle for a CaP expert system capable of simulating urology experts for managing patients with stable prostate cancer in any clinical setting.
- To validate the CaP expert system, complying with the industrial and MHRA standards.

1.6 Contribution to knowledge

There are three main knowledge gaps analysed in this thesis. First, the gap between ES production and implementation in health care. This was systematically analysed by establishing state of the art in Urology then compare them against standards from the industry. The second gap is the lack of formality of ES development and their compliance with the regulatory bodies' standard. Third gap of knowledge acquisition was analysed by demonstrating diverse industrial models and implementing a novel model for acquiring knowledge from Urologist.

1.7 Thesis organisation

Chapter two evaluates state of the art of ES systems in Urology. They were systematically reviewed following the PRISMA methodology demonstrating their availability, development and application in Urology. Further evidence gleaned on their development were applied ES development in this thesis.

Chapter three presents the literature review of CaP demographics, treatment modalities and biochemical failure definitions. Those are the core parameters applied by a Urologist in CaP follow up decision making and were used to as the factual knowledge for developing our ES knowledgebase.

Chapter four describes the CaP procedural knowledge acquisition study where real data was used to elicit knowledge from human experts. The outcome was represented in decision trees for further validation.

Chapter five includes the validation of the hand written decision trees by a novel mixed methodology validation model developed for this study. The validated knowledge was then integrated in the ES for developing the follow up computer based follow up.

Chapter six describes and reports the multicentre validation of this follow up model by comparing the system recommendation to two independent Urology experts.

The last chapter of this thesis, Chapter seven, summarises the work and the potential of those systems in health care. It also highlights the possible directions of future research in this field.

Chapter two - Systematic review of Artificial intelligence in Urology

2.1 Introduction

This study aimed was to systematically review published ES in Urological health care with a primary aim to demonstrate their availability, progression, testing and applications. The secondary aim was to evaluate their development life cycle against standards suggested by Okeefe and Oleary (1993) and Benbasat and Dhaliwal (1989) in their review articles on ES development. The later would evaluate the gap between their development and implementation in health care.

2.2 Methods

Data for this review was gathered following the PRISMA methodology. For the primary aim, data was collected and analysed on system model, year of production, type and outcome of its validation, functional domain application, variables for input and output, target user and domain. This analysis demonstrated their prevalence, testing, and applications.

Further qualitative analysis was required to meet the study secondary aim. For this, further data was gathered on credibility (user perception on the system) , evaluation (system usability), validation (building the right system) and verification (building the system right) then compare against the standards reported in Okeefe and Oleary (1993) and Benbasat and Dhaliwal (1989).

Eligibility criteria included any published article reporting an ES developed to simulate experts in urological domains. We considered all published systems from 1960 until September 2016 as we did not expect the ES technology to be available or mature before 1960 (Liao, 2005).

Inclusion criteria consisted of systems following the definition of ES and applied in Urological domains. The study should clearly demonstrate the system model and describe its components and development. To formulate our search, we used the models described in Liao et al. 10 years review of the literature (Rule based, Knowledge based, Neural networks, Fuzzy systems, Object-oriented, Case based reasoning, System architecture development, Intelligent agent systems, Modelling ontology and Database methodology) to identify eligible articles from their key words (Liao, 2005).

Further models' category (not reported in Liao et al.'s review) included hybrid systems integrating intelligent and knowledge computing models as fuzzy logic - neural network system where one system exhibits both models. These were described in in Pandey et al. review article on medical ES development (Pandey and Mishra, 2009) where ES were categorised according to the ratio between the data to human knowledge utilised in the system development.

We excluded any system which was not computerised such as paper based tools or software based on simple algorithms instead of the known models of ES as reported Liao et al. (Liao, 2005) and Pandey et al. (Pandey and Mishra, 2009).

Furthermore, we excluded any ES that were only applied in medical education or for developing other systems or were used as intelligent data base or as part of medical equipment. Other systems were excluded for insufficient information (failed quality assessment) on the system to satisfy the inclusion. Thus, we excluded published abstracts and review articles. Upon including more than one hundred articles, the rest of the eligible articles were compared to the ones included, then excluded based on demonstrating clear similarity. This was applied to avoid expanding the size of the data without adding to the study analysis.

The information sources considered in our study are WEB OF SCIENCE, EMBASE, BIOSIS CITATION INDEX, SCOPUS, and MEDLINE.

Our search protocol included Urology as topic AND ES common keywords (Table 2.1 WEB OF SCIENCE literature review protocol)

Study selection was performed through examining the title and abstract review to identify eligible articles. A full text review was then performed to examine the study against the inclusion criteria.

Quality assessment and data gathered for each system included: model, author, year of publication, application, input & output variables, validation strategy, study population (data or live – how many centres), statistical outcome and target user. Articles failed the quality assessment if missing any of these data with exception of statistical outcome, validation strategy and target users. Further systems were excluded in the quality assessment when demonstrating similarity to included systems and did not report a new model of development or application or validation method. Thus, not adding to the study analysis and results.

The data were collected following a full text review of all included articles by the lead author and documented on an electronic spread sheet (Excel, Microsoft Corp, Seattle).

Risk of bias was not performed as the outcome of the individual system testing was not a main objective of our systematic review nor synthesised in the data analysis.

The systems were further stratified in the quantitative analysis according to our interpretation of its core application in Urology in one or more of the seven groups if the system was designed to perform more than one function. These seven groups were: Decision Support (DS), Diagnosis Prediction (DP), Disease Stage prediction (DSP),

Treatment Outcome prediction (TOP), Predicting Research Variables (PRV) and Image Analysis (IA).

#1	**TOPIC:** ("expert system*") OR **TOPIC:** ("decision support") OR **TOPIC:** ("artificial intelligence") OR **TOPIC:** ("rule based") OR **TOPIC:** ("knowledge base* system*") OR **TOPIC:** ("neural network") OR **TOPIC:** ("fuzzy") DocType=All document types; Language=All languages;
#2	**TOPIC:** (urology) DocType=All document types; Language=All languages;
#3	#1 AND #2 DocType=All document types; Language=All languages;

Table 2.1 WEB OF SCIENCE literature review protocol
Common Expert Systems models (rule based, neural networks, fuzzy logic) were applied in the search to identify compatible systems in combination to Urology as the main topic. This identified eligible articles reporting Expert systems applied in Urological domains.

2.3 Results

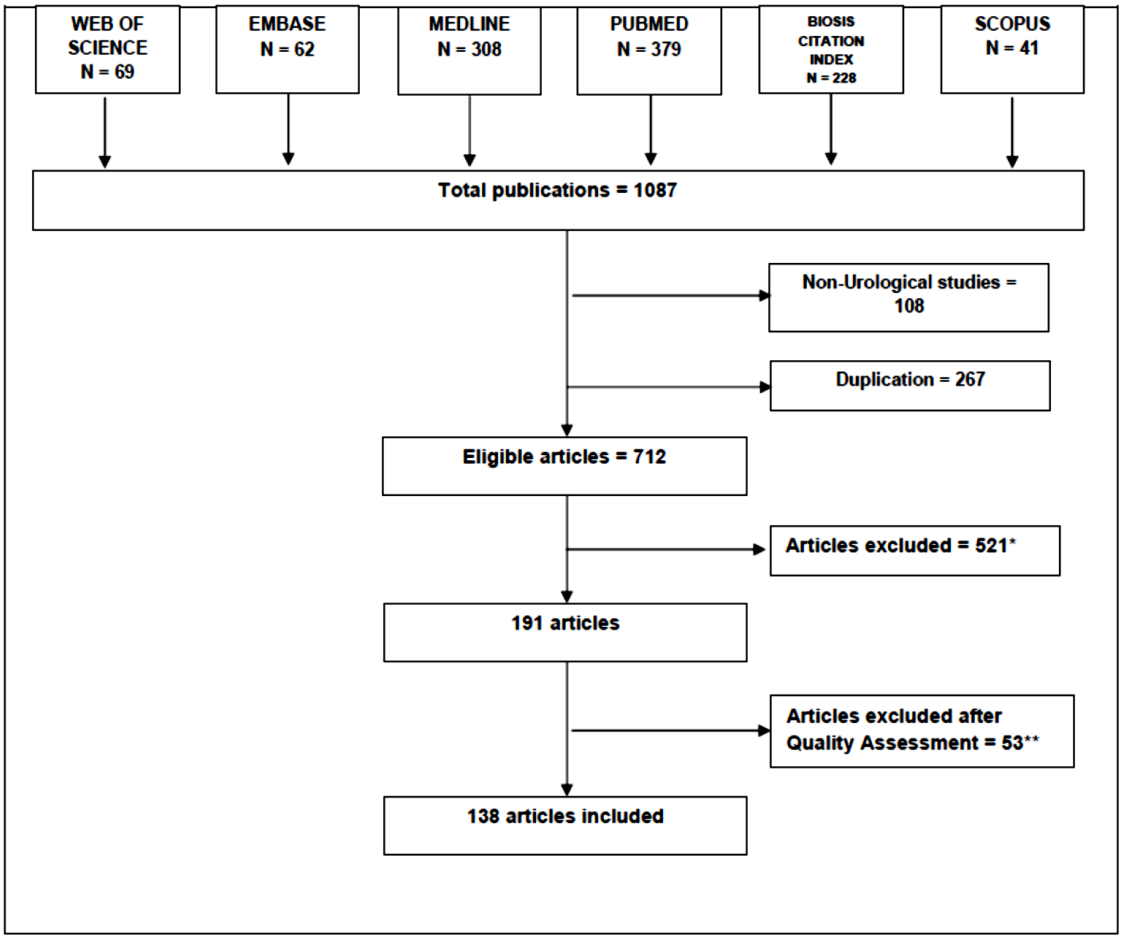


Figure 2.1 PRISMA flow chart for the systematic review of articles included in the review of expert systems in Urology

**explanation below in excluded articles table*

***articles were not included on quality assessment mainly due to core similarity to included articles and few were missing important information on system development (thus excluded)*

Reasons	Total
Review articles	95
Editorials / Letters	19
Research does not include Expert System	374
Expert system for design of medical devices	10
Expert systems for medical education	5
Non-human study	1
Conference abstract	19
Total	521

Table 2.2 Excluded articles

Articles were excluded when the tool structure was not consistent with the known models and functions of expert systems. The number of the review articles was amplified by the duplicated articles and non-urological reviews.

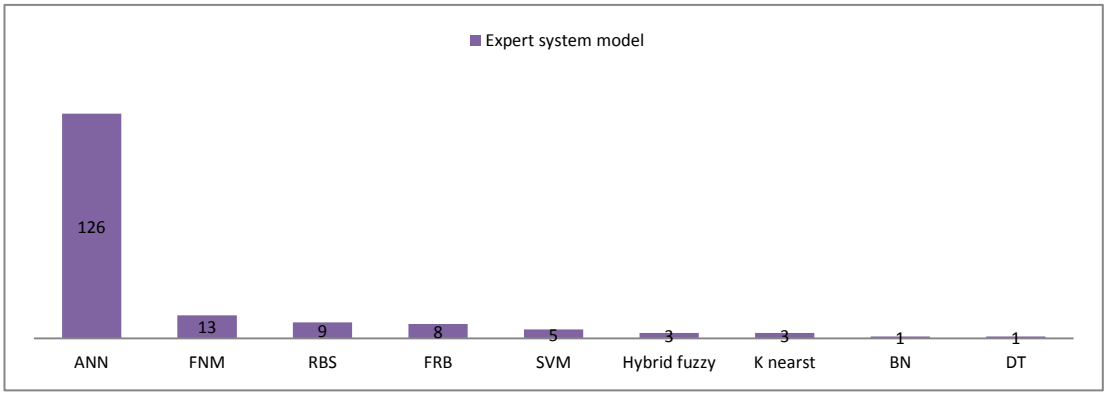


Figure 2.2 Analysis of Expert systems by models.
ANN was the most common but other systems were applied on different domain as fuzzy neural model (FNM), rule-based system (RBS), fuzzy rule based (FRB), support vector machine (SVT), Bayesian network (BN) and decision trees (DT).

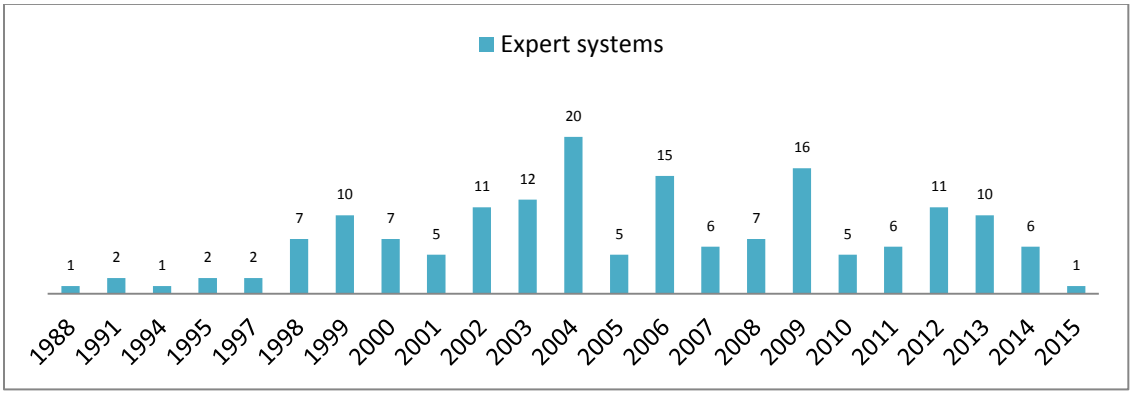


Figure 2.3 Expert system analysis by year of publication demonstrating an upward trend and increase in number of publications.
Systems were included according to the keywords for expert system models and applied in urological domains

ANN was the commonest model to be applied in Urological ES (Figure 2.2). The rest of the models demonstrated diversity which is consistent with other published industrial systems (Liao, 2005).

Prostate cancer was the commonest domain for urological ES (Figure 2.4) with most of the system focusing on cancer diagnosis. These systems were applied to various domains (Figure 2.4) and they were further stratified and analysed according to their core functional application as outlines in the methodology.

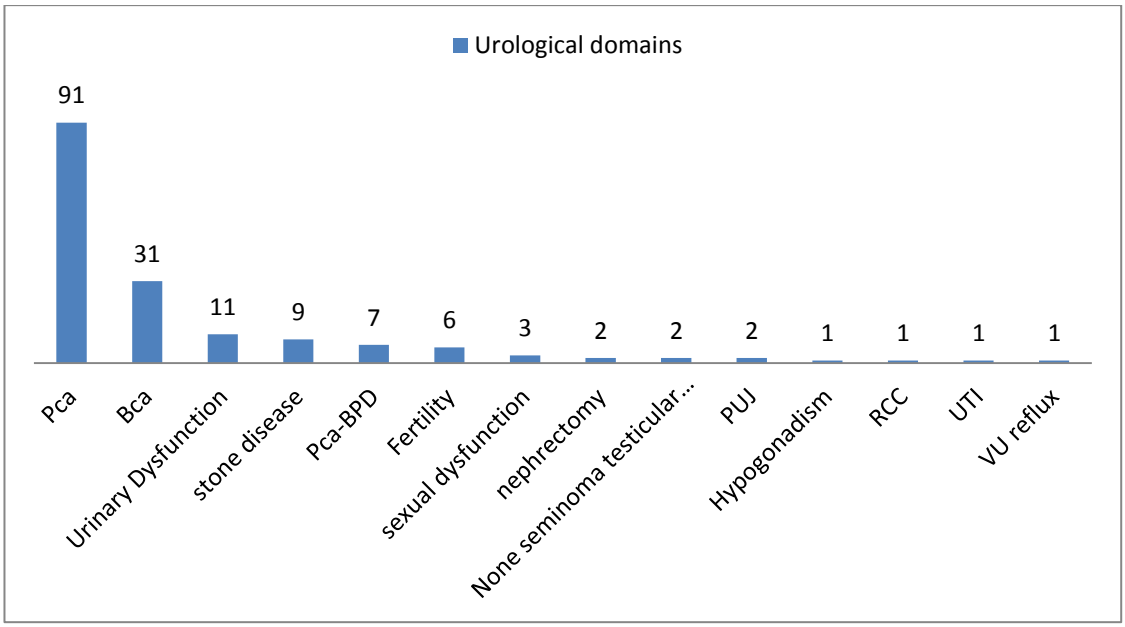


Figure 2.4 Urological domains applied by Expert systems. Prostate cancer (CaP) was the commonest domain followed by bladder cancer (Bca) then other diseases as benign prostatic disease (BPD), pelvi ureteric junction obstruction (PUJ), Urinary tract infection (UTI), renal cell cancer (RCC), vesico ureteric reflux (VU reflux).

2.3.1 **Quantitative analysis**

2.3.1.1 Decision Support systems

The main objective of ES in this domain is to facilitate the clinical decision making by identifying key elements from patients clinical and laboratory examinations then refine a theoretical diagnostic or treatment strategy (Koutsojannis et al., 2009). They can guide the expert to find the right answer (Petrovic et al. (2011)) or take over the decision making to support the none expert as (Keles et al. (2007) or even replace both to interact with the patient directly (Gorman (1995).

They have supported various aspects of urological decision making such as diagnosis, investigations analysis, radiotherapy dose calculation, the delivery of behavioural treatment and therapeutic dialogues.

Table 2.3 Decision support systems in Urological domain

A total of 21expert systems included supporting the decision making in Urological domains. Rule based reasoning was the most common model and urinary dysfunction was the commonest domain

Article	Mdl	Dom	Subdomain	variables	Output	Knowledge acquisition	Validation method	Target user
(Petrucci et al., 1991)	RBR	UDys	Incontinence in long-term care facilities	Disease related questions	Recommendations	Experts	Comparison to blinded experts & pilot RCT	Non-expert nurses
(Gorman, 1995)	RBR	UDys	U incont treatment	Incontinence symptoms	Behavioural treatment	Agency guidelines	RCT (60) reliability and validity by experts	Patients
(Boyington et al., 2004)	RBR	UDys	U incont treatment	19 evaluation questionnaires	Individualised health information	An expert & patients' feedback	No validation	Patients
(Koutsojannis et al., 2012)	RBR	UDys	U incont	MH, incontinence symptoms, previous incidents and medication history	U incont treatment	Multiple experts, patients record and literature	Evaluation by experts, 95 retrospective data	Non-experts
(Hao et al., 2013)	RBR	UDys	Ward management of micturition	LUTS, Urinary tract infection Anatomical obstruction, Multiple causality and sensory impairment	Diagnosis and risk of fall	Multiple experts	Se 0.95, Sp 0.72, Likert scale Cronbach α 0.9	Urology ward nurses
(Sucevic and Ilic, 1991)	FRB	UDys	U dyn interpretation	U dyn variables	Detrusor and sphincter dysfunction	Not mentioned	Improve User Ac by 10%	Experts
(Altunay et al., 2009)	ANN	UDys	Uroflow interpretation	Value of slopes, frequency and value of maximums, ration of amplitude and total voiding time	Healthy or pathologic Uroflow	Patients data from U dyn	78 test cases ROC 0.7 Ac 79%	Experts
(Gil and Johnsson, 2010)	SVM	UDys	Diagnosis	Age, examination, Uroflow, U dyn	Healthy or pathologic Uroflow	Patients data	Ac 84%, Se 93%, Sp 33%	Experts
(Lopes et al., 2013)	FNM	UDys	Diagnosis	46 defining Characteristics from NANDA-I	Diagnosis of U Dys	Multiple experts weighted the variables and literature review	kappa vs experts (0.92-0.42), Se 0.95, Sp 0.92	Experts & non-experts

(Keles et al., 2007)	FNM	Ca P-BPD	Diagnosis of BPE and CaP	Clinical and pathological variables	CaP, BPE medical, BPE surgery	Patients data	10 folds CV AUC 0.86, se 100%, sp 98%	Non-experts
(Koutsojannis et al., 2008)	FRB	Ca P-BPD	AP CP CaP BPE	LUTS, quality of life, fever, haematuria, haemospermia, painful ejaculation, fever, perineal pain, bone pain, pyuria, age, DRE.	Diagnosis and treatment of prostatic disease	Multiple experts interviews, patients records and literature	Ac 0.76, Se 0.79, Sp 0.75, retrospective data (n=105)	Residents, patients, medical students
(Koutsojannis et al., 2009)	FRB	Ca P-BPD	AP CP CaP BPE	LUTS, quality of life, fever, haematuria, haemospermia, painful ejaculation, fever, perineal pain, bone pain, pyuria, age, DRE.	Diagnosis and treatment	WEKA* to extract rules then experts to modify	200 test cases Ac 0.93, Se 0.97, Sp 0.99,	Residents, patients, medical students
(Pereira et al., 2004)	RBR	CaP	Diagnosis before 1st biopsy	Age, race, FH, DRE, PSA, PSAD, PSAV, TRUS findings	Cancer and benign	Not mentioned	25 test cases Se 100% Sp 33% PPV 62%, NPV 100%	Experts
(Petrovic et al., 2011)	F-CBR	CaP	Radiotherapy dose for CaP	Gl, PSA, Distribution Volume Histogram	Radiotherapy dose	72 patients' cases	Comparison to experts, Ac 85%	Experts
(Torshizi et al., 2014)	F-ONT	BPD	Diagnosis and treatment of BPE	LUTS, DRE	Watchful waiting, medical, surgery	Multiple experts weighted the variables	44 prospective cases, agreement kappa = 0.89	Experts & non-experts
(Binik et al., 1988)	RBR	SDys	Diagnosis and treatment	Set of descriptors	Therapeutic dialogue	Not mentioned	10 Patients' evaluations	Couples
(Beligiannis et al., 2006)	RBR	SDys	Male S dys diagnosis	22 parameters from history and examination	ED diagnosis	GA rule extraction from 30 cases	Se (73-94%), Sp (78-96%) Ac (89%) vs Residents	Un specified

(Koutsojannis and Hatzilygeroudis, 2004)	FRB	SDys	Male S dys diagnosis & treatment	MH, non-coital erection, diabetes mellitus, coronary artery, neuropathies, sexual history, psychosocial history, depression, smoking, alcohol, examination, hormonal evaluation, cholesterol	Diagnosis and treatment of ED	Multiple experts' interviews, Pearson analysis on variables from patients' data and literature	70 test cases vs experts and non-experts (Ac79%)	Non-experts
(Papageorgiou, 2012)	FNM	UTI	UTI treatment	Clinical data on UTI	Antibiotics course	Patients data & guidelines	Ac 86.8 %, 38 random cases	Experts & non-experts
(Arlen et al.)	ANN	VUR	Decision support for intervention	Age, gender, number of UTIs prior to VUR diagnosis, UTI, of complete ureteral duplication noted on Ultrasound, the presence of bowel or bladder dysfunction.	UTI or not	255 cases, 96 cases	AUC 0.76	Experts
(Goyal et al., 2007)	ANN	Nlt	ESWL dose calculation	Age, stone size, stone burden, number of sittings	Number and power of shock	196 cases, 80 cases	coefficient of correlation 0.9	Experts

.

2.3.1.1.1 Domains

Urinary dysfunction (U Dys) was the commonest domain to be covered in the decision support system application (n=9), which could be further categorised into U Dys diagnostic, investigation analysis and therapeutic systems. They have demonstrated a range of methodologies, validation and target users (Table 2.3 Decision support systems in Urological domain). For instance, Hao et al. (2013) designed an ES to support junior nurses in diagnosing urinary elimination dysfunction in a selected group of patients while Lopes et al. (2013), Koutsojannis et al. (2012) systems were able to support any medical user to diagnose urinary incontinence with an accuracy reaching higher than 90%. The target user of most of these systems were predominantly medical health care workers including both experts and none experts, with exception of Gorman (1995) Boyington et al. (2004) which can be directly used by patients to receive an assessment of their urinary elimination dysfunction followed by a tailored treatment plan.

Urodynamic studies are the most informative and diagnostic investigation in this domain where the bladder pressure figures are recorded at the filling and voiding phases. The interpretation of these figures is quite complex even to the Urologist where they must receive specialist training before, they can accurately report it. This reporting was remarkably facilitated by an ES developed by Sucevic and Ilic (1991) where it improved the expert's accuracy by 10%. Uroflow meter is another useful bed side test in this domain that can estimate the efficacy of voiding by plotting the volume against the speed on a graph. The figures and patterns of this graph can be processed by Urologists to diagnose a number voiding pathologies. Altunay et al. (2009) ES was designed to identify the normal flow study with an overall accuracy of 79%. Uroflow and

urodynamic diagnostic power were both fed into Gil and Johnsson (2010) 'system which had an overall sensitivity of 93% for detecting urinary dysfunction.

Prostate diseases were represented in 6 systems while 3 of them modelled by Koutsojannis et al. (2008), Koutsojannis et al. (2009), Keles et al. (2007) for diagnosing both benign and malignant prostatic disease.

All systems in this domain were diagnosis support system with exception of Torshizi et al. (2014) which also provided treatment for benign prostatic hyperplasia (BPH) and Petrovic et al. (2011) calculated the required radiotherapy dose for treating CaP.

Sexual dysfunctions were modelled in 3 systems where Beligiannis et al. (2006) diagnosed male sexual dysfunction with an accuracy of 89%, while Koutsojannis and Hatzilygeroudis (2004) added a therapeutic model for the same disease with an overall accuracy of 79%. Sexpert by Binik et al. (1988) was the third system in this category developed in 1988 and in fact the oldest ES to be identified from our search in all urological domains. Interestingly this RB system was designed to interact directly with couples suffering from sexual dysfunction where the system responds to their query with a tailored therapeutic dialogue for treating their problem.

Urinary tract infection (UTI) was diagnosed and treated by one of the hybrid fuzzy systems FNM developed by Papageorgiou (2012) with an accuracy of 86.8%.

2.3.1.2 Diagnosis prediction

In this domain, ES aim to aid the practitioner in reaching a diagnosis by analysing clinical variables and quantifying the probability of a clinical diagnosis with a defined margin of error. This is usually presented to the clinician as an estimated probability of the diagnosis tailored for each individual case. They simulate the opinion of a second

expert or act as a useful but non-invasive bed side test. It has been suggested that their use could eliminate unnecessary invasive investigation as suggested by Remzi et al. (2003) that their ANN could reduce up to 68% of unnecessary repeated TRUS biopsies to diagnose CaP.

Table 2.4 Diagnosis prediction application of Expert systems in Urology

A total of 37 systems identified in this application of expert systems in Urology with evident prevalence of ANN as the model and CaP to be the dominant domain.

Article	Mdl	Dom	Subdomain	Variables	Output	System training	Validation	Statistical outcome
(Ronco and Fernandez, 1999)	ANN	CaP	Pre-biopsy diagnosis with TRUS variables	Age, PSA, number of biopsies, clinical diagnosis, PSAD, TRUS variables	Cancer or benign	N = 442 from single centre database	ROC AUC NPV, PPV ½ CV	NPV 97%, PPV 82% better than LR
(Babaian et al., 2000)	ANN	CaP	Diagnosis PSA 2.5-4	Age, tPSA, creatinine phospho kinase, prostatic acid phosphatase	Cancer or benign	Multicentre data 522 (PSA 2.5-4)	ROC AUC CV 152 cases	AUC 0.74
(Finne et al., 2000)	ANN	CaP	Diagnosis PSA 4-10	Age, tPSA, %fPSA, TPV, DRE	Risk of Cancer	656 data from Finnish trial	ROC, Sp, Se LOO	Se 79%, Sp 57%, Ac 62%, PPV 35, NPV 90
(Stephan et al., 2002)	ANN	CaP	Diagnosis PSA 2-20	Age, tPSA, %fPSA, TPV, DRE	Risk of Cancer	1188 multi centre	ROC, Sp, Se, 1/10 CV	Sp 90%, Se 64%
(Stephan et al., 2002)	ANN	CaP	Diagnosis in trial patients with PSA 4-10	Age, tPSA, %fPSA, TPV, DRE	Risk of Cancer	1188 multi centre	ROC, Sp, Se, 204 trial data PSA 4-10	Se 95%, Sp 23.3%, CI 17.4%–30.2%, P < 0.0002
(Djavan et al., 2002)	ANN	CaP	Diagnosis	fPSA, TZD, PSAV, %f PSA, TZV, t PSA, and PSAD	Cancer or benign	PSA 2.5-4, 272 patients, multicentre data	ROC, AUC ¼ CV	AUC 0.88
(Djavan et al., 2002)	ANN	CaP	Diagnosis	TZD, % f PSA, PSAD and TPV	Cancer or benign	PSA 4-10, 974 patients, multicentre data	ROC, AUC ¼ CV	AUC 0.91
(Remzi et al., 2003)	ANN	CaP	Diagnosis after initial negative biopsy PSA 4-10	t PSA, %f PSA, TPV, TZV, PSAD, TZD	Cancer or benign	820 patients with PSA 4-10 European cancer detection studies	ROC AUC 1/3 CV	AUC 0.83
(Kalra et al., 2003)	ANN	CaP	Diagnosis of BPE and CaP	Age, ethnicity, FH, IPSS, t PSA, %f PSA, DRE	Risk of cancer	Multicentre 354 patients, multicentre	ROC vs LR, 144 test set 40% CV	AUC ANN 0.8, LR 0.5
(Saritas et al., 2003)	FRB	CaP	Early diagnosis	Age, t PSA, TPV	Risk of cancer	Experts aided in developing 77 fuzzy rules	Not published	None
(Matsui et al., 2004)	ANN	CaP	Diagnosis PSA 2-10	Age, tPSA, %fPSA, TPV, TZV, PSAD, TZD = ANNA 1	Cancer and benign	228 data one centre	ROC, 30% CV	AUC 0.78

(Matsui et al., 2004)	ANN	CaP	Diagnosis PSA 2-10	ANNA 1 + presumed circle area ratio and DRE	Cancer and benign	228 data one centre	ROC 30% CV	AUC 0.79, Sp 45%, Se 90%
(Porter et al., 2005)	ANN	CaP	Diagnosis	Age, tPSA, TPV, PSAD, DRE, and TRUS findings	Cancer and benign	3814 prostate cancer screening data	ROC AUC 1/3 CV, 2 centres prospective data	AUC: 0.74, 0.76, and 0.75 prospective 0.73, 0.74
(Lee et al., 2006)	ANN	CaP	Diagnosis	Age, DRE, PSA, PSAD, TZV, TZD = ANNA	Cancer and benign	TRUS, single centre 684 data	ROC AUC 1/4 CV	AUC 0.74
(Lee et al., 2006)	ANN	CaP	Diagnosis	ANNA + TRUS findings	Cancer and benign	TRUS, single centre 684 data	ROC AUC 1/4 CV	AUC 0.86
(Benecchi, 2006)	FNM	CaP	Diagnosis, PSA < 20	Age, PSA, %f PSA	Cancer and benign	1030 patients' data, one centre	ROC, Sp, Se, 1/4 CV	AUC 0.8, Sp 52%, Se 90%
(Stephan et al., 2007)	ANN	CaP	Prostate cancer early diagnosis PSA 4-10	Age, tPSA, %fPSA, TPV, DRE	Cancer or benign	606 multicentre group (PSA 4-10)	ROC AUC, 1/10 CV	AUC 0.83, AUC 0.74 in Finish group
(Stephan et al., 2007)	ANN	CaP	Prostate cancer early diagnosis PSA 4-10	Age, tPSA, %fPSA, TPV, DRE	Cancer or benign	656 Finnish cancer survey group (PSA 4-10)	ROC AUC, 1/10 CV	AUC 0.77
(Kawakami et al., 2008)	ANN	CaP	Diagnosis	Age, DRE, t PSA and f PSA	Cancer and benign	1509 with PSA < 20, Single centre	ROC AUC, 1/5 CV	AUC 0.74
(Kawakami et al., 2008)	ANN	CaP	Diagnosis	Age, DRE, t PSA, f PSA, TPV and TRUS findings	Cancer and benign	1509 with PSA < 20, Single centre	ROC AUC, 1/5 CV	AUC 0.75
(Stephan et al., 2009)	ANN	CaP	Diagnosis with -2 Pro PSA	Age, TPV, tPSA, %fPSA, p2 PSA, %p2 PSA (-2 proPSA)	Cancer and benign	PSA 1-30, 586 one centre	ROC, Sp, Se LOO 586	AUC 0.85, Sp 62%, Se 90%
(Lee et al., 2010)	ANN	CaP	Diagnosis pre-biopsy	Age, DRE, tPSA, PSAD, TZD, TRUS findings	Benign and malignant	600 patients with suspected CaP	ROC AUC, 477 random	AUC 0.77
(Lee et al., 2010)	SVM	CaP	Diagnosis pre-biopsy	Age, DRE, tPSA, PSAD, TZD, TRUS findings	Benign and malignant	600 patients with suspected CaP	ROC AUC, 477 random	AUC 0.85

(Meijer et al., 2009)	ANN	CaP	Diagnosis PSA 2-20	Age, tPSA, %f PSA, DRE, TPV	Cancer and benign	Testing Prostataclass	ROC AUC, 165 patients one centre	AUC (PSA 2-10) 63-69%, (PSA 10-20) 57-88%
(Saritas et al., 2010)	ANN	CaP	Diagnosis	Age, tPSA, %f PSA	Prognosis: cancer or not	121 Patients data from one centre	ROC AUC, 30 % CV 29 patients	AUC 0.92
(Lawrentschuk et al., 2011)	ANN	CaP	Diagnosis of clinically significant cancer	Age, DRE, PSA, PRV, TRUS, Biopsy cores	Disease clinical significance	3025 multicentre data	Accuracy estimation	Ac 57%
(Ecke et al., 2012)	ANN	CaP	Diagnosis of cancer	Age, DRE, PSA, %fPSA, and TPV	Cancer and benign	204 PSA between 4 -10	ROC AUC	AUC 0.72
(Filella et al., 2014)	ANN	CaP	PHI index and TPV in diagnosis	Age, TPV, %fPSA, tPSA, PHI, %P2PSA	Cancer and benign	220 cases PSA < 10	ROC AUC	AUC 0.81
(Filella et al., 2014)	ANN	CaP	PHI index and TPV in diagnosis	Age, %fPSA, tPSA, PHI, %P2PSA	Cancer and benign	221 cases PSA < 10	ROC AUC	AUC 0.77
(Yuksel et al., 2013)	FRB	CaP	Diagnosis	Age, PSA, TPV	Cancer and benign	78 TRUS cases from Urology clinic	None	None
(Samli and Dogan, 2004)	ANN	Fert	Sperm count	Age, duration of infertility, FSH, LH, TT and PRL, testicular volume	Presence of spermatozoa	303 patient's data	ROC AUC then kappa stats of LR, test set 73 random	Se 68%, Sp 87.5%, PPV 73.9%, NPV 84%
(Powell et al., 2007)	ANN	Fert	Endocrinopathy with low sperm count	Testis volume, total sperm count,	Endocrinopathy	1035 Data from 2 centres	ROC AUC	AUC 0.95
(Ramasamy et al., 2013)	ANN	Fert	Microdissection testicular sperm extraction	Age, FSH level, cryptorchidism and Klinefelter syndrome	Sperm retrieval	1026 data, one centre	ROC AUC	Se 67% Sp 49.5% PPV, 63.9% NPV 52% Ac 60.8%
(Paya et al., 2006)	ANN	UDys	Interpretation of U dyn and symptoms	Neurological & physical symptoms, flowmetry, cystometry, U dyn	Areflexia, hyper-reflexive, effort incontinence	400 U Dyn data	80 patients, 1/5 CV,	Accuracy 85%
(Gil et al., 2009)	ANN	UDys	Interpretation of U dyn and symptoms	Neurological & physical symptoms, flowmetry, cystometry, U dynamics	Healthy or ill	300 patients with LUT disease	ROC, Ac, 1/5 CV	Accuracy 89%

(Wadie et al., 2006)	ANN	UDys	Bladder outlet obstruction	values of the average flow rate, Qmax, PVR and TPV	Obstructed, Non-obstructed, and equivocal	N = 457 cases from single centre	Accuracy estimation 157 cases	Ac 60% (testing) 75% (training)
(Wadie et al., 2001)	ANN	BPD	IPSS interpretation	IPSS subdomain scores	Obstructed, Non-obstructed, and equivocal	N = 460 from single centre	Accuracy estimation 157 cases	Ac 73%

2.3.1.2.1 Domains

Prostate cancer diagnosis had a strong presence in this domain with 26 out of the 29 systems identified for application. One of the earliest systems was developed by Babaian et al. (2000) with variables of a serum based 'ProstAsure index' designed by Stamey et al. (1996) to improve CaP detection in patients with BPH. This ProstAsure (age, total PSA, creatinine phosphokinase isoenzymes and prostatic acid phosphatase) based ANN was tested on patients 'cases with BPH and low PSA < 4ng/ml and reported a good sensitivity of 93% and specificity of 81% in diagnosing CaP before TRUS guided biopsy.

Later systems have introduced other variables to their network inputs such as digital rectal examination findings (DRE), % free (free/total) PSA (%f PSA), total prostate volume (TPV), PSA density (PSAd), transitional zone volume (TZV) and transitional zone PSA density (TZD) with the %fPSA being the commonest variable among all system inputs (Table 2.4). More TRUS based variables were investigated by Lee et al. (2006) & Matsui et al. (2004) where TRUS +ve findings improved the prediction accuracy of Matsui et al. (2004) network ($p<0.05$, AUC 86%). Similarly, the PCAR (presumed circle area ratio: is the ratio of the area of the maximum horizontal section of the prostate by TRUS to the area of a presumed circle of which the circumference is equal to the circumference of the maximum horizontal section) increased the system prediction (AUC 83% $p<0.05$) of Lee et al. (2006) when added to other PSA variables.

Other researchers added further BPH variables as the International Prostate Symptom Score (IPSS) which was modeled in neUROn system by Kalra et al. (2003). This model has reported improvement of Pca prediction in a multicentre analysis of patients' data

with suspected cancer (abnormal DRE and/or PSA>4) by adding IPSS to the variables when compared to the linear regression (AUC 83%, p<0.002).

The target population of these systems are patients with suspected Pca which is usually based on elevated PSA levels. Eight systems in this domain targeted the cohort with low or borderline elevated levels (PSA<10) including two reported by Babaian et al. (1998), Djavan et al. (2002) predicting cancer in patients with PSA<4 (AUC 87%, p<0.02 for %f PSA).

The implementation of these systems in the prostate cancer diagnosis domain is expected to exclude unnecessary biopsy Finne et al. (2000), Remzi et al. (2003) where systems can reduce unnecessary biopsies by 13% and 68% respectively.

Urinary dysfunction (U Dys) ES was applied by Paya et al. (2006) and Gil et al. (2009) to interpret urodynamic results and patients' symptoms to predict the diagnosis with an overall accuracy reaching 89%.

Fertility ES Samli and Dogan (2004) predicted the presence of spermatozoa in the sperm analysis by combining variables on clinical features, testicular volume and hormonal level.

2.3.1.3 Disease stage prediction

Disease staging is a process of disease risk analysis, aiming to stratify the patients into graded risk groups in order to estimate the disease progression and plan treatment. In cancer, this ensures optimal treatment and avoid over or under treating the condition. This process is usually achieved through collaborative effort of cancer multidisciplinary team analysing available clinical, radiological and histological information. Despite their best effort, there is always a margin of error in the assessment including the

limitation of the available clinical and imaging tools. An example of this limitation is the group of patients who had a clinical and radiological diagnosis of localised prostate cancer and considered candidates for curative surgery. It is known that a number of these patients will have evidence of advanced disease in their surgical specimen, which was not picked up in the routine staging and they can suffer from long term poor prognosis (Han et al., 2001a). The main aim of the systems in this application is to act as a staging tool to provide accurate disease staging even before obtaining the definitive operative histology.

Table 2.5 Disease stage prediction

This table demonstrated expert systems predicting urological diagnosis from variable clinical and radiological date. Artificial neural networks (ANN) diagnosing localised prostate cancer (Pca) before surgery were the most common systems in this application.

Art	Mdl	Dom	Subdomain	Variables	Output	System training	Statistical outcome	Validation set
(Tewari and Narayan, 1998)	ANN	CaP	staging of localised disease	Age, race, DRE, tPSA, size of tumour on ultrasound, Gl, bilaterality of cancer and number of positive cores and perineural infiltration	Margin, seminal vesicle and lymph node positivity	1200, patients' data from multicentre	AUC 0.77, 0.79, 0.8	20% CV
(Chang et al., 1999)	FSS	CaP	Localised disease staging	Age, PSA, PSAD, DRE, TRUS, Gl, CT, bone scan, chest x-ray, MRI	Localised or advanced	16 Cases	Se 92%, Sp 84%, Ac 82%	43 cases RRP
(Batuello et al., 2001)	ANN	CaP	Lymph node staging in CaP post RPP	Age, Gl, clinical stage	Lymph node spread	736 data from one centre clinically localised CaP	Se 64%, Sp 81.5%, PPV 14%, NPV 98%	1840 & 316 cases from 2 centres
(Han et al., 2001b)	ANN	CaP	Prostate cancer staging post RRP	Age, tPSA, Gl, clinical stage	Lymph node spread or organ confinement	5744 data from one centre clinically localised CaP	AUC 77%, 88% for LN	25% CV random
(Mattfeldt et al., 2001)	ANN	CaP	Stage prediction post RRP	Age, histological variables from biopsy	CaP stage	97 cases with non-organ confined	Prediction accuracy ranged from 82% to 90%	
(Mattfeldt et al., 2001)	ANN	CaP	Stage prediction post RRP	Age, histological variables from biopsy, tPSA and TPV	CaP stage	77 cases with non-organ confined and extracapsular spread	Prediction accuracy ranged from 82% to 90%	
(Matsui et al., 2002)	ANN	CaP	Prostate cancer staging post RRP PSA 2-10	tPSA, TNM, Gl (ANNA1)	localised disease	124 data from 2 centres Clinically localised CaP	AUC 0.82	20 % (n=36 patients)
(Matsui et al., 2002)	ANN	CaP	Prostate cancer staging post RRP PSA 2-10	tPSA, TNM, Gl, maximum tumour length (ANNA2)	localised disease	124 data from 2 centres Clinically localised CaP	AUC 0.88	20 % (n=36 patients)
(Matsui et al., 2002)	ANN	CaP	Prostate cancer staging post RRP PSA 2-10	tPSA, TNM, Gl, maximum tumour length, PSAD (ANNA3)	localised disease	124 data 2 centres Clinically localised CaP	Ac 83.3%, Se 85%, Sp 83%,	20 % (n=36 patients)

							PPV 73%, NPV 90% AUC 0.9	
(Matsui et al., 2002)	ANN	CaP	Prostate cancer staging post RRP PSA 2-10	tPSA, TNM, Gl, maximum tumour length PSAD, age (ANNA4)	localised disease	124 data 2 centres Clinically localised CaP	AUC 0.87	20 % 36 patients
(Zlotta et al., 2003)	ANN	CaP	Prostate cancer staging post RRP	tPSA, TPV, TZV, PSAD, TZ, Gl	Pathological stage t2-4	201 cases from multinational European cancer data base (PSA 10 or less)	AUC 0.87	61 prospective set
(Chiu et al., 2009)	ANN	CaP	diagnosis of skeletal metastasis	Age, tPSA	skeletal Mets	111 retrospective cases in one centre	AUC 0.88, Se 87.5%, Sp 83.3%	Bootstrap CV
(Kim et al., 2011)	ANN	CaP	Stage prediction post RRP	DRE, % of cancer, sum of tumour length, % cancer length & maximum cancer core length	advanced cancer (> pT3a)	300 randomly selected from retrospective data	AUC 0.71, Se 63%, Sp 81%, Ac78%	232 random selected set
(Kim et al., 2011)	SVM	CaP	Stage prediction post RRP	DRE, % of cancer, sum of tumour length, % cancer length & maximum cancer core length	advanced caner (> pT3a)	300 randomly selected from retrospective data	AUC 0.81, Se 67%, Sp 79%, Ac77%	232 random selected set
(Regnier-Coudert et al., 2012)	ANN	CaP	Define precise stage	PSA, clinical stage, pathological stage, Gl (other added for different set: erection, IPSS, TRUS size, MRI stage	margin, seminal vesicle and lymph node positivity	From 7500 patients' data from BAUS database and remodelled with external data of 85 patients	AUC 0.38-0.67, concordance index for variables	10 folds CV
(Regnier-Coudert et al., 2012)	BN	CaP	Define precise stage	PSA, clinical stage, pathological stage, Gl (other added for different set: erection, IPSS, TRUS size, MRI stage	margin, seminal vesicle and lymph node positivity	From 7500 patients' data from BAUS database and remodelled with external data of 85 patients	AUC 0.01-0.67 concordance index for variables	10 folds CV
(Regnier-Coudert et al., 2012)	kNN	CaP	Define precise stage	PSA, clinical stage, pathological stage, Gl (other added for different set: erection, IPSS, TRUS size, MRI stage	margin, seminal vesicle and lymph node positivity	From 7500 patients' data from BAUS database and remodelled with external data of 85 patients	AUC 0.33-0.6 concordance index for variables	10 folds CV
(Regnier-Coudert et al., 2012)	RBF	CaP	Define precise stage	PSA, clinical stage, pathological stage, Gl (other added for different set: erection, IPSS, TRUS size, MRI stage	margin, seminal vesicle and lymph node positivity	From 7500 patients' data from BAUS database and remodelled with external data of 85 patients	AUC 0.45-0.5 concordance index for variables	10 folds CV
(Regnier-Coudert et al., 2012)	SVM	CaP	Define precise stage	PSA, clinical stage, pathological stage, Gl (other added for different set: erection, IPSS, TRUS size, MRI stage	margin, seminal vesicle and lymph node positivity	From 7500 patients' data from BAUS database and remodelled with external data of 85 patients	AUC 0.5 concordance index for variables	10 folds CV

(Veltri et al., 2002)	ANN	CaP	Staging post RRP	Age, tPSA, n Positive cores, involvement per core, % of positive core	Organ confinement and metastasis	870 multicentre data	Ac 60%	120 cases, Accuracy estimation
(Cosma et al., 2016)	FNM	CaP	Cancer staging of organ confinement	Age, PSA, Primary Gleason Pattern, secondary Gleason pattern, clinical stage	Organ confinement and metastasis	399 cases from research network database	AUC 0.8, FNM outperformed ANN, FCM, LR	ROC AUC vs other models
(Moul et al., 1995)	ANN	Nsc	staging	vascular, lymphatic, tunical invasion, percentage of embryonal carcinoma, yolk sac carcinoma, teratoma and seminoma	Stage one or two	93 cancer specimen, single centre	Prediction accuracy 79.6 to 87.1%,	10 folds CV

2.3.1.3.1 Domain

Prostate cancer was the main domain for this application with 19 systems out of 20. Most of them were designed to predict organ confinement before radical surgical excision of the prostate (Table 2.5). The target population were patients with clinically localised CaP and their accuracy reached high estimates as in (Batuello et al., 2001), where the system was able to predict 98% of the low risk group for lymph node involvement using preoperative available date (PSA, clinical stage and Gleason score).

Chiu et al. (Chiu et al., 2009) Modelled a system with clinical variables for patients undergoing nuclear bone scintigraphy for predicting skeletal metastasis. The system was able to predict metastatic disease in the test group with Se 87.5%, Sp 83.3%.

None seminoma testicular cancer was the other domain in this application with the system (Moul et al., 1995) able to predict the cancer disease stage with accuracy reaching 87%.

2.3.1.4 Treatment outcome prediction

Deciding the appropriate treatment instigates weighing the benefits of the intervention against its known risks while considering the patient's perspective.

The treatment benefits including its success rate are influenced by multiple confounding factors, which by default are none linear or standard for each individual case. Further case analysis will rely on the clinicians' experience. ES aim to enhance this process by quantifying this multivariate relation and estimating the outcome of treatment for each individual case. As in (Poulakis et al., 2002, Hamid et al., 2003, Gomha et al., 2004, Michaels et al., 1998) where the system predicted the outcome of nephrolithiasis extra

corporeal shock wave treatment (ESWL) and (Porter et al., 2001, Potter et al., 1999) providing an estimation of biochemical failure (BCF) after radical prostatectomy.

Table 2.6 Treatment outcome prediction

Is one of the common applications of urological expert system. They predicted treatment outcome of radical nephrectomy, radical cystectomy, radical prostatectomy, vesico ureteric reflux endoscopic repair, pelvi-ureteric junction obstruction conservative management, nephrolithiasis conservative management and extracorporeal shockwave treatment. The commonest domain was predicting negative surgical margins post radical prostatectomy.

Art	Mdl	Dom	Subdomain	Variables	Output	System training	Validation Methods	Statistical Outcome
(Naguib et al., 1998)	ANN	CaP	Outcome of RRP	Age, stage, bone scan, grade, PSA, treatment, bcl-2, p54	No response, response then relapse, response & no relapse.	cohort of CaP single centre 21 patients	ROC, Sp, Se 20 patients randomly selected	Ac 85% (60% without markers), K, 0.65; Cl, P<0.00001
(Potter et al., 1999)	ANN	CaP	BCF post RRP	Age Pathologic findings & GENN1	Disease progression	Gl 5-7, T1B-2C, Single centre 136	ROC, Sp, Se Test set of 35 (20 %)	AUC 0.71, Ac 74%, Se 82%, Sp 61%,
(Potter et al., 1999)	ANN	CaP	BCF post RRP	DNA polyploidy and quantitative nuclear grade	Disease progression	Gl 5-7, T1B-2C, Single centre 136	ROC, Sp, Se	AUC 0.74, Ac 80%, Se 75%, Sp 85%
(Potter et al., 1999)	ANN	CaP	BCF post RRP	Pathologic findings, age, DNA polyploidy & quantitative nuclear grade	Disease progression	Gl 5-7, T1B-2C, Single centre 136	Test set of 35 (20 %)	AUC 0.73 Ac 78%, Se 84%, Sp 72%
(Porter et al., 2001)	ANN	CaP	BCF post RRP	Age, PSA, Gl & stage	BCF post RRP all	140 cases post RRP, one centre	ROC, Sp, Se 35 (20%) for validity	AUC 0.81, Se 74%, Sp 78%, PPV 71%, NPV 81%,
(Seker et al., 2003)	Fkn	CaP	Outcome of RRP	TM, Gl, PSA, P53, bcl-2, treatment method	No response No progression after treatment, Relapse	41 men with CaP	LOO and compare predictive accuracy of ANN, Fkn	Predictive accuracy ranged from 61-88%
(Zlotta et al., 2003)	ANN	CaP	Outcome of RRP	tPSA, TZV, PSAd, Gl	Local or advanced disease	200 cases from multinational European cancer data base	AUC ROC 60 prospective set	AUC 0.91, Se 95%, Sp 64%,

(Poulakis et al., 2004a)	ANN	CaP	Outcome of RRP, margin positive	tPSA, clinical stage, Gl (ANNA1)	Positive surgical margins	218 post RRP and pelvic lymph adenectomy in one centre	ROC AUC 48 cases 1/4 CV	AUC 0.7
(Poulakis et al., 2004a)	ANN	CaP	Outcome of RRP, margin positive	tPSA, clinical stage, Gl, pMRI findings (ANNA2)	Positive surgical margins	218 post RRRP and pelvic lymph adenectomy in one centre	ROC AUC 48 cases 1/4 CV	AUC 0.87
(Poulakis et al., 2004a)	ANN	CaP	Outcome of RRP, margin positive	tPSA, clinical stage, Gl, pMRI findings, % of cancer in biopsy, PSAd ANNA3	Positive surgical margins	218 post RRP and pelvic lymph adenopathy in one centre	ROC AUC 48 cases 1/4 CV	AUC 0.87
(Poulakis et al., 2004a)	ANN	CaP	Outcome of RRP, margin positive	tPSA, clinical stage, Gl, % of cancer in biopsy ANNA4	Positive surgical margins	218 post RRP and pelvic lymph adenopathy in one centre	ROC AUC 48 cases 1/4 CV	AUC 0.71
(Poulakis et al., 2004b)	ANN	CaP	Outcome of RRP, margin & LN	tPSA, clinical TNM Gl ANNA1	Positive surgical margins, LN involvement	41 post RRP and pelvic lymph adenopathy in one centre	ROC AUC 160 cases randomly selected	AUC 0.86 for positive margin, 0.88 for LN +ve
(Poulakis et al., 2004b)	ANN	CaP	Outcome of RRP, margin & LN	tPSA, clinical TNM Gl, pMRI findings ANNA2	Positive surgical margins, LN involvement	41 post RRP and pelvic lymph adenopathy in one centre	ROC AUC 160 cases randomly selected	AUC 0.9 for positive margin, 0.89 for LN +ve
(Poulakis et al., 2004b)	ANN	CaP	Outcome of RRP, margin & LN	tPSA, clinical stage, Gl, pMRI findings, age ANNA3	Positive surgical margins, LN involvement	41 post RRP and pelvic lymph adenopathy in one centre	ROC AUC 160 cases randomly selected	AUC 0.9 for positive margin, 0.9 for LN +ve
(de Paula Castanho et al., 2008)	FRB	CaP	Outcome of RRP	Clinical stage, Gl, tPSA	Cancer stage (confined, capsule, vesicle & LN)	116 rules developed from nomograms	ROC Se, Sp 190 patients post RRP in one centre	AUC 0.76 (95% CI 0.7-0.8), Se 85%, Sp 61%)
(Botoca et al., 2009)	ANN	CaP	Outcome of RRP, margin positive	TNM stage, age, Gl, tPSA,	Capsule penetration	650 retrospective data for RRP at one centre	PPV, NPV 98 cases for testing & 1/2 CV	PPV 100%, NPV 95%

(Botoca et al., 2009)	ANN	CaP	Outcome of RRP, margin positive	TNM stage, age, Gl, tPSA MLP,	Capsule penetration	650 retrospective data for RRP at one centre	PPV, NPV 98 cases for testing & 1/2 CV	PPV 97%, NPV 95%
(Botoca et al., 2009)	ANN	CaP	Outcome of RRP, margin positive	TNM stage, age, Gl, tPSA, Partial RNN (recurrent neural network)	Capsule penetration	650 retrospective data for RRP at one centre	PPV, NPV 98 cases for testing & 1/2 CV	PPV 97%, NPV 95%
(Botoca et al., 2009)	ANN	CaP	Outcome of RRP, margin positive	TNM stage, age, Gl, tPSA, RBF-MLP	Capsule penetration	650 retrospective data for RRP at one centre	PPV, NPV 98 cases for testing & 1/2 CV	PPV 97%, NPV 94%
(Castanho et al., 2013)	FRB	CaP	Outcome of RPP	Clinical stage, Gl, tPSA	Capsule penetration	Genetic algorithm on 331 patients post RRP in one centre	48 patients post RRP in one centre ROC	AUC 0.82 (95% CI 0.5-0.8)
(Hu et al., 2014)	ANN	CaP	Outcome of LAP RRP, BCF	Clinical and pathologic parameters, tPSA, margin status, TNM and Gl	BCF	1575 patients at one centre post lap RRPP	ROC AUC LOO	AUC 0.75, Se 90%, Sp 35
(Goyal et al., 2007)	FNM	CaP	Outcome post RRP	Age, FH, DRE, tPSA, Gl, MR findings	tPSA at 6 months	19 one centre post RRP	Correlation coefficient = 0.99	3 Cases
(Tewari et al., 2001)	ANN	CaP	Outcome post RRP	Age, tPSA, staging, perineural infiltration, Gl, months of FU	BCF	1400 multicentre data	Se 85% Sp74%, PPV 77%	400 data
(Borque et al., 2001)	ANN	CaP	Outcome post RRP, organ confined	Gleason score, preoperative PSA and clinical stage,	Organ confined	468 cases for training	NPV 83%	47 cases 30% CV
(Tsao et al., 2014)	ANN	CaP	Outcome of RRPP	PSA, BMI, DRE, TRUS, Gl score or grade	Capsule penetration	225 patients' data post RRP from 3 centres	74 patients randomly selected ROC	AUC 0.79 LR 0.74 (P = 0.016) Partin AUC 0.7
(Michaels et al., 1998)	ANN	Nlt	Stone regrowth after ESWL	Anatomy, position, stone analysis, urine analysis, previous stone, medical treatment	Stone recurrence	single centre data base, 65 cases	ROC, Sp, Se33 cases	Se 91%, Sp 92%, AUC 0.96
(Cummings et al., 2000)	ANN	Nlt	Stone clearance with conservative treatment	Age, gender, duration, creatinine, nausea, vomiting, fever	Clearance or intervention	multi centre, Ureteric stone 125 cases	55 cases ROC, Sp, Se	AC 76% Predict 100% of stones passed

(Poulakis et al., 2002)	ANN	Nlt	lower pole stone ESWL	Gender, BMI, radiology, stone size & composition, urine analysis, 24 hrs urine, serum ca & creatinine	Clearance or intervention	321 patients with lower pole stone	211 random set ROC, Sp, Se, vs LR	AUC 0.97 Se 95%, Sp 92%,
(Hamid et al., 2003)	ANN	Nlt	Stone clearance with ESWL	Age, gender, body habitus, serum electrolytes, 24 hrs urine, radiological findings	Stone free	60 patients, one centre	Correlation co-efficient 22 cases	0.75
(Gomha et al., 2004)	ANN	Nlt	Stone clearance with ESWL	Age, gender, anatomy, location, side, number, length, width, new or recurrent, stent	Stone clearance	Ureteric stone ESWL, One centre 688 cases	296 cases ROC, Sp, Se	Ac 78%, Se78%, Sp 75%, PPV 97%
(Dal Moro et al., 2006)	ANN	Nlt	Outcome of conservative stone disease treatment	Age, gender, BMI, fever, previous treatments & stones, duration of the symptoms, dimension and position of the stone	Spontaneous expulsion or intervention	402 patients from one centre	50 patient, 1/4 cross validation ROC Se, Sp	Se 95%, Sp 63%
(Dal Moro et al., 2006)	SVM	Nlt	Outcome of conservative stone disease treatment	Age, gender, BMI, fever, previous treatments & stones, duration of the symptoms, dimension and position of the stone	Spontaneous expulsion or intervention	402 patients from one centre	50 patient, 1/4 cross validation ROC Se, Sp	Se 85%, Sp 87%
(Sun and Chang, 2006)	ANN	Nlt	ESWL outcome prediction	The patients' characteristics, stone location, burden, shape dimension, pre-ESWL procedure and cost of admission	unexpected post-ESWL visits	1026 patients received ESWL at one centre`	AUC 0.66	506 patients
(Bagli et al., 1998)	ANN	PUJ	Outcome of PUJ repair	Demographic, clinical and radiological findings	Sonographic outcome of pyeloplasty	Single centre unilateral paediatric pyeloplasty n=100	16 cases (16%) ROC, Sp, Se	Ac 100%, Se 100%, Sp 100%
(Seçkiner et al., 2011)	ANN	PUJ	Outcome of PUJ conservative treatment	Age, gender, renal pelvis diameter, laterality, separated renal function on DMSA, urine culture and infections	Observation or surgery	37 infants with PUJ obstruction	Prediction accuracy16 patients for validation	75 % prediction accuracy
(Parekattil et al., 2005)	ANN	Neph	Post lap partial nephrectomy hospital stay	Age, co-morbidities, tumour size and extension	Hospital stay less than 2 days	334 one centre	5 institutes 77, 19 prospective ROC	AUC 0.6, 0.5
(Parekattil et al., 2005)	ANN	Ne	Post lap nephrectomy hospital stay	Age, co-morbidities, tumour size and extension	Hospital stay less than 2 days	392 One centre	5 institutes 127, 29 prospective ROC	AUC 0.7, 0.7

		ph						
(Vukicevic et al., 2014)Z	ANN	Bca	Pathological stage after surgery	Age, gender, tumour (size, number, grade, invasion, lymph vascular invasion, stage), lymph nodes	Prognosis and advanced stage	183 patients, one centre post cystectomy	ROC & compare with LR 1/3 cross validation	MANN AUC 0.86, Se 88%, Sp 77%, PPV 93%, NPV 63%, Ac 85%
(Vukicevic et al., 2014)	ANN	Bca	Pathological stage after surgery	Age, gender, tumour (size, number, grade, invasion, lymph vascular invasion, stage), lymph nodes	Prognosis and advanced stage	183 patients, one centre post cystectomy	ROC & compare with LR 1/3 cross validation	SANN AUC 0.85, Se 84%, Sp 71%, PPV 91%, NPV 67%, Ac 83%
(Serrano-Durba et al., 2004)	ANN	VUR	outcome of endo repair of VU reflux	Age, gender, implant type, implant volume, number of treatments, side, endo findings, type of cystography	Ultrasound finding	Single centre data base, paediatric VU reflux 174 data	87 cases for validation ROC, Sp, Se	Se 71.4%, Sp 81.6%, PPV 58.8%, NPV 88.6% and success rate 78.9%,

2.3.1.4.1 Domains

Prostate cancer was also common domain in this application (n=23). Potter (Potter et al. (1999). Porter et al. (2001)) described 4 models developed by data acquired from patients with clinically localised CaP and had radical prostatectomy with curative intent. The variables included clinical and histological findings of the surgical specimen and they were able to predict up to 81 % who did not have evidence biochemical failure (rising PSA) in their follow up. Seker et al. (2003) and Naguib et al. (1998) models were not restricted to the clinically localised CaP cohort and their study population included patients at different disease stages and on any treatment pathway. Their models included 2 experimental histological markers (tumour suppressor gene p53 and the proto-oncogene bcl-2) in their input variables and the estimated predictive accuracy of the patient response to treatment were reaching 68% and 80% ($p<0.00001$) respectively.

Nephrolithiasis treatment was expressed by 6 other systems applying the treatment outcome prediction concept. It is a common cause of emergency hospital admission while only a small proportion requiring an intervention to relieve their symptoms or clear the stone burden. *Cummings et al.* targeted this group in his ANN (Cummings et al., 2000) where he trained his network with patients' data treated at the emergency service of 3 centres with ureteric stones, to identify patients failing conservative management and requiring further intervention. When tested on a different set of 55 cases, the system correctly predicted 100% of the patients who passed the stone spontaneously with an overall accuracy of 76%.

Extra corporeal shockwave lithotripsy (ESWL) is one of the favourable interventions in the nephrolithiasis treatment domain. The stone here receives strong external shock waves, which can subsequently reduce it into small fragment and eliminate the need for

direct instrumentation of the renal tract. Their reported success rate can only provide a generalised prediction of outcome to the individual case and ANN was capable of providing an alternative multivariate analytical tool in the 4 models developed by Poulakis et al. (2002) Hamid et al. (2003) Gomha et al. (2004) and Michaels et al. (1998). They estimated high accuracy of their models (Table 2.6), as in (Gomha et al., 2004), the system predicted 97% of the patients who were confirmed to be stone free following ESWL for treating ureteric stone.

Paediatric pelvi-ureteric junction obstruction is primarily treated conservatively unless there is any evidence of renal function compromise, recurring infection or worsening radiological findings. For the failing group, pyeloplasty is the second line of treatment and Bagli et al. (1998) developed an ANN to estimate the success rate of this procedure for each individual case by predicting the post-operative degree of hydronephrosis with a reported 100% accuracy in the small tested sample. Seçkiner et al. (2011) modelled their system to predict the outcome of managing the disease conservative and the system demonstrated a predictive accuracy of 75%.

Vesico ureteric reflux or reflux uropathy is another paediatric disease, characterised by back flow of urine from the bladder into the ureter through incompetent Vesico ureteric functional valve. It can present with recurrent UTI and their treatment is primarily conservative as it can be a self-limiting disease. The second option is surgical reimplantation of the ureters or endoscopic injection of bulking agent at the ureteric orifices. Serrano-Durba et al. (2004) trained a neural network using 261 cases whom have received endoscopic injection and the system predicted 94% of the patients who did not benefit from the treatment.

Laparoscopic partial and radical nephrectomy were the domain of the Parekattil et al. (2005), which was developed by multi institutional case data (age, co-morbidities, tumour size, and extension) of patients having laparoscopic partial or radical nephrectomy. The system was able to predict the length of their postoperative hospital stay with an accuracy of 72%.

Bladder cancer can be treated with complete bladder excision and Vukicevic et al. (2014) developed systems to predict success of this procedure and positive margins in the surgical specimen with an accuracy of 83%.

2.3.1.5 Recurrence and survival prediction:

The ES in this domain aimed to provide individual risk analysis tools estimating the disease specific mortality and recognising the group whom may benefit from more aggressive or adjuvant treatment.

Table 2.7 Recurrence and progression prediction

The majority of the expert systems in this application were artificial neural network predicting recurrence and survival following bladder cancer treatment. Other systems were applied in non-seminomatous testicular cancer, prostate cancer, renal cell carcinoma and recurrence of upper renal tract stones.

Art	Mdl	Dom	Subdomain	Variables	Output	Knowledge acquisition	Validation	Statistical outcome
(Naguib et al., 1997)	ANN	Bca	Recurrence	Age, gender, smoking, tumour stage and grade, CIS, number, cytology, other mucosal biopsy	Recurrence or no	N = 432 patients' data, multicentre	Radom set of 200 ROC AUC	Se 76%, Sp 55%, Ac 72%
(Qureshi et al., 2000)	ANN	Bca	Tumour progression recurrence	Tumour stage and grade, size, number, gender, eGFR	Stage progression	105 Ta/T1 TCC multicentre	Compare to 4 clinicians McNemar test	80% accuracy
(Qureshi et al., 2000)	ANN	Bca	12 months cancer specific survival	Tumour stage and grade, size, number, gender, eGFR, smoking, cis, dysplasia tumour site, architecture, c-erbB2 (oncogene), p53 (tumour suppressor gene)	6 months recurrence 12 months survival	56 Ta/T1 (6 months recurrence), 40 T2-T4 (12 months survival)	Compare to 4 clinicians McNemar test	Accuracy to predict recurrence (75%) and to predict survival (82%)
(Fujikawa et al., 2003)	ANN	Bca	Progression of non-invasive TCC	Age, gender, tumour (grade, stage, number & architecture) and mean nuclear volume	Tumour progression & recurrence	68 patients' specimen from one centre	22 Random test set ROC, Sp, Se	Recurrence: Se 33%, Sp 40%, PPV 40%, NPV 33% Progression: Se 100%, Sp 67%, PPV 40%, NPV 100%
(Catto et al., 2003)	FNM	Bca	Recurrence classifier	Age, gender, grade, smoking, previous cancer, p53, hMLH1, hMSH2	Recurrence or not	109 patients from one centre with TCC	10% cross validation ROC, LR	AUC 0.98, Se 90%, Sp 80%, PPV 92%, NPV 74%, Ac 88%
(Catto et al., 2003)	FNM	Bca	Survival predictor	Age, gender, grade, smoking, previous cancer, p53, hMLH1, hMSH2	Survival in months	109 patients from one centre with TCC	10% cross validation Root mean square	RMS = 4.8
(Catto et al., 2003)	ANN	Bca	Recurrence classifier	Age, gender, grade, smoking, previous cancer, p53, hMLH1, hMSH2	Recurrence or not	109 patients from one centre with TCC	ROC, LR 10% cross validation	AUC 0.91, Se 94%, Sp 96%, PPV 99%, NPV 84%, Ac 95%
(Catto et al., 2003)	ANN	Bca	Survival predictor	Age, gender, grade, smoking, previous cancer, p53, hMLH1, hMSH2	Survival in months	109 patients from one centre with bladder	10% cross validation RMS	RMS = 11.7

(ABBOD et al., 2004)	ANN	Bca	Survival predictor	Age, stage, Grade, smoking, previous cancer	Risk of relapse	109 patients with primary TCC	Difference in RMS 1/4 CV ROC AUC	Se 90%, Sp 89%, PPV 98, NPV, 64% Ac 90%, RMS 8.8
(ABBOD et al., 2004)	ANN	Bca	Recurrence predictor	Stage, Grade, age, smoking, previous cancer, p53, hMLH1, hMLH2	Time to relapse	109 patients with primary TCC	Difference in RMS 1/4 CV ROC AUC	Se 94, Sp 96%, NPV 99%,PPV 84%, Ac 95%, RMS 7.6
(ABBOD et al., 2004)	FNM	Bca	Survival predictor	Stage, Grade, age, smoking, previous cancer	Risk of relapse	109 patients with primary TCC	Difference in RMS 1/4 CV ROC AUC	Se 92%, Sp 90%, PPV 98% NPV 72%, Ac 92%, RMS 8.5
(ABBOD et al., 2004)	FNM	Bca	Recurrence predictor	Stage, Grade, age, smoking, previous cancer, p53, hMLH1, hMLH2	Time to relapse	109 patients with primary TCC	Difference in RMS 1/4 CV ROC AUC	Se 90% Sp 80%, NPV 92%,PPV 74%, Ac 88%, RMS 7.3
(Catto et al., 2006)	FNM	Bca	Recurrence (classifier)	Age, gender, grade, smoking, previous cancer, p53, methylation index (% of loci on chromosomes), RARB methylation (gene locus)	Recurrence or not	117 patients with 1ry TCC or UCC from one centre	10% cross validation ROC, LR	AUC 0.98, Se 88-100%, Sp 94-100%, Ac 100%
(Catto et al., 2006)	FNM	Bca	Survival predictor	Age, gender, grade, smoking, previous cancer, p53, methylation index (% of loci on chromosomes), RARB methylation	Survival in months	117 patients with 1ry TCC or UCC from one centre	10% CV Kaplan Maier for survival	Average error = 5 months
(Catto et al., 2006)	ANN	Bca	Recurrence (classifier)	Age, gender, grade, smoking, previous cancer, p53, methylation index (% of loci on chromosomes), RARB methylation	Recurrence or not	117 patients with 1ry TCC or UCC from one centre	10% cross validation ROC, LR	Ac 89-90%, Se 81-87%, Sp 95-100%
(Catto et al., 2006)	ANN	Bca	Survival predictor	Age, gender, grade, smoking, previous cancer, p53, methylation index (% of loci on chromosomes), RARB methylation	Survival in months	117 patients with 1ry TCC or UCC from one centre	10% CV Kaplan Maier for survival	Average error = 9 months
(Cai et al., 2007a)	ANN	Bca	Recurrence	Age, sex, previous recurrence, response to adjuvant therapy, number of lesions, adjuvant therapy	Recurrence or no	403 patients	1/3 CV 123 patients ROC AUC	AUC 0.87, Se 79%, Sp 98%
(Bassi et al., 2007)	ANN	Bca	5 Years survival cystectomy	Age, gender, tumour stage, grade, ln, vascular in, perineural in, prostatic invasion, CaP	Survival at 5 years	369 patients	ROC, Cox proportional hazard 1/3 CV	Se 63%, Sp 86%, PPV 76%, NPV 77%
(Catto et al., 2009)	FNM	Bca	Recurrence classifier	Gender, pathological stage, grade, CIS, lymph vascular invasion	Recurrence or not	609 patients from multiple centres	ROC, LR 10% CV	Se 93 %, Sp 68 %

(Catto et al., 2009)	FNM	Bca	Survival predictor	Gender, pathological stage, grade, CIS, lymph vascular invasion	Survival in months	172 multicentre data	ROC, LR 10% CV	Kaplan-Meier survival plots, median error of 8.15 months
(El-Mekresh et al., 2009)	ANN	Bca	Survival post cystectomy	Age, gender, bilhariziasis, histology, grade, lymph nodes, lymph vascular, type of diversion	Patient survival	871 patients' data post cystectomy	30% CV ROC vs LR	AUC 0.86, Se 79%, Sp 81%
(Kolasa et al., 2009)	ANN	Bca	bladder cancer 5 years survival	Age, gender, histology grade, tumour stage, positive LN, removed LN	5 years survival	cystectomy data base, single centre 106 patients	Prediction error percent 11 and 29 patients	prediction error rate, >90% efficiency
(Buchner et al., 2013)	ANN	Bca	Recurrence & survival	Age, gender, tumour stage, grade, CIS, ln, lymph vascular invasion	5 years recurrence & cancer specific death	cystectomy data base, multicentre 2111	ROC, Kaplan Maier for survival, Cox Proportional Hazard	Se 59%, Sp 77%, PPV 67%, NPV 70% (30% cross validation)
(Wang et al., 2015)	ANN	Bca	Survival post cystectomy	Age, gender, albumin, surgical approach, tumour stage, follow up period, type of diversion	5 years survival	117 patients with post cystectomy from one centre	10 Folds CV ROC, Se, Sp Ac	Ac 72-80% RELM & ELM had best performance
(Cai et al., 2007b)	ANN	Bca	Recurrence of G3 pTa after TURBT	Age, sex, previous histopathological data, previous recurrence rate response to previous BCG adjuvant therapy, number of lesions, size of lesions presence of inflammatory reaction and adjuvant therapy.	Recurrence or No	143 patients with G3 pTa at one centre	AUC, Se, Se 1/3 cv 43 cases	AUC 0.81, Se 82%, Sp 96%
(Buchner et al., 2012)	ANN	RCC	RCC survival 36 months	Age, gender, BMI, performance status, histopathology, time interval between primary tumour and detection of Mets, type of systemic therapy, number & sites of Met	Recurrence within 36 months	175 single centre	30 % CV ROC sensitivity analysis	AUC 0.95 (95% CI 0.87-0.98)
(Marszall et al., 2012)	ANN	Nsc	Disease recurrence in five years	(32 variables) age, tumour type, grade, invasion, Mets, ln, treatment, FBC, kidney function	Recurrence within five years	202 multicentre cases	1/4 CV ROC, Sensitivity analysis	AUC 0.87
(Kuo et al., 2015)	FNM	CaP	Prognosis and survival	Age, BMI, PSA, DRE, Gl, clinical stage and treatment methods.	Disease specific survival	100 cases single centre	Sensitivity analysis mean square error,	MSE = 0.068907 (1/10 CV)
(Tanthanuch and Tanthanuch, 2004)	ANN	Nlt	Recurrence of Upper tract stones	Age, sex, history of previous calculi, radiologic type, location & composition of previous calculi, 24-hour urine assay urine culture, treatment	Recurrence of Upper tract stones	168 cases, single centre	PPV estimation 68 cases	PPV 100%

2.3.1.5.1 Domains

Bladder cancer survival and recurrence prediction following radical cystectomy (RC) with curative intention was the commonest domain in this application (24 out of 26 total systems). It is the 4th commonest cause of cancer and the 6th leading cause of cancer related death in the male population of the UK in 2011 (UK, 2011). Among all cancers, it has the highest recurrence rate with more than 40% recurrence post RC for muscle invasive tumours (W. Herr et al., 2007, von der Maase et al., 2005). The nodal involvement is highly predictive of the recurrence and these patients are considered for adjuvant or neoadjuvant systemic chemotherapy. The node free cohort will include some high-risk patients for disease progression and mortality who were not identified by the conventional linear stratification system. *Catto et al.* developed a FNM system to identify this high risk group in the nodal free cohort by predicting the disease recurrence rate (Se 81%, Sp 85%) and their survival with a median error of 8.15 months (Catto et al., 2009). The high-risk group identified by this model can benefit from systemic treatment post cystectomy to improve their disease related morbidity and mortality (von der Maase et al., 2005, W. Herr et al., 2007). The 5 years survival post cystectomy was the output of 2 other ANN Bassi et al. (2007) Kolasa et al. (2009) with the inputs including facts on patients' demographics, tumour invasion and positive lymph nodes in the surgical specimen. Bassi et al. (2007) predicted 77% of the disease specific mortality in 5 years and Kolasa et al. (2009) experimental model reached prediction efficacy of > 90%.

Renal cell cancer is primarily treated with partial or radical nephrectomy for clinically localised disease with adjuvant systemic therapy for the metastatic disease. There is still a degree of uncertainty in stratifying individual disease risk in order to predict the

indication and outcome of systemic therapy in the group with distant metastasis. Buchner et al. (2012) attempted to clarify this uncertainty by training a neural network with data patients' data who had nephrectomy (partial or radical) and received systemic therapy. The mature model predicted the patients who survived the disease at 3 years with an overall accuracy of 95% (CI 0.878-0.987).

None seminoma testicular cancer 5 years recurrence was the domain of Marszall et al. (2012) ANN. The system was trained with multicentre data and in its testing phase it predicted 100% of the patients who did not suffer from disease recurrence at 5 years with an overall predictive accuracy of 94% (AUC=87%).

2.3.1.6 Predicting research variables

In academia, testing a hypothesis for 'factors-outcome effect' is a popular quest and the standard statistical regression analysis tools may not be effective for data contaminated by too many noisy variables (Catto et al., 2010). AI can provide an alternative methodology in analysing these data to identify variables with high correlation to the outcome by applying machine learning as in ANN.

A common methodology estimates the receiver operating characteristics (ROC) and calculates the area under the curve (AUC) for the whole system after the process of machine learning is completed. Then research variables are randomised among all cases to see if this would affect the AUC Stephan et al. (2009). When the AUC decreases this can be interpreted that this variable is significant and other variables can be tested using the exact methodology. In effect, the wider the discrepancy of the AUC the more significant the variable is.

Table 2.8 Research variable prediction

In this application, the system modifies their machine learning ability to identify the significant variables from the data in terms of their correlation to a specified outcome. This can save time, effort and cost specially when applied on gene microarrays

Art	Mdl	DOM	Subdomain	Variables	Output	System training	Validation	Statistics	Research outcome
(Krongrad et al., 1997)	ANN	BPE/CaP	Analysis of variables of quality of life questionnaire	Questionnaire suggested by medical and allied professional	High- or low-quality group	Single centre recruitment with BPE or CaP, 63 cases	ROC, Linear quadratic and logistic regression	Ac 90%, Se 94%, Sp 85%, PPV 89%, NPV 92%	Identify relevant variables
(Michaels et al., 1998)	ANN	Nlt	Stone recurrence after ESWL	Anatomy, position, stone analysis, urine analysis, previous stone, medical treatment	Stone recurrence	65 patients post ESWL from single centre	33 test set ROC AUC vs LR	AUC 0.96, Se 91%, S 91%	Stone recurrence, fragments not risk factor
(Han et al., 2000)	ANN	CaP	Biochemical failure post RRP	TNM, tPSA, Gleason, pathology stage	BCF at 3 years Yes or no	564 patients' data post RRRP with Gl 7, single centre	ROC, Kaplan Meier and Cox Proportional Hazards Model	AUC 75%, NPV 84	Gleason 7 is inversely correlated to disease free survival and direct to BCF
(Han et al., 2000)	ANN	CaP	Biochemical failure post RRP	TNM, tPSA, Gleason, pathology stage	BCF post RRRP	564 patients' data post RRRP with clinically localised CaP Gl7, single centre	ROC, Kaplan Meier for survival and Cox Proportional Hazards	AUC 81%, NPV 93%	
(Poulakis et al., 2002)	ANN	Nlt	lower pole stone ESWL	Gender, BMI, radiology, stone size, composition, urine analysis, 24 hrs urine, serum ca & creatinine	Clearance or intervention	321 patients with lower pole stone	211 random set ROC, Sp, Se, vs LR	AUC 0.97 Se 95%, Sp 92%	BMI, normal urinary transport & infundibular width of 5 mm or more and the infundibular ureteropelvic angle is 45° or more are correlated with stone clearance
(Catto et al., 2003)	FNM	Bca	Recurrence classifier	Age, gender, grade, smoking, previous cancer, p53, hMLH1, hMSH2	Recurrence or not	109 patients from one centre with bladder TCC	10-fold CV ROC, LR	AUC 0.98, Se 90%, Sp 80%, PPV 92%, NPV 74%, Ac 88%	p value calculated to compare all models, the effect of combining HK p53 with other variables
(Catto et al., 2003)	FNM	Bca	Survival predictor	Age, gender, grade, smoking, previous cancer, p53, hMLH1, hMSH2	Survival in months	109 patients from one centre with bladder TCC	10-fold CV Root mean square	RMS = 4.8	

(Catto et al., 2003)	ANN	Bca	Recurrence classifier	Age, gender, grade, smoking, previous cancer, p53, hMLH1, hMSH2	Recurrence or not	109 patients from one centre with bladder TCC	ROC, LR 10% cross validation	AUC 0.91, Se 94%, Sp 96%, PPV 99%, NPV 84%, Ac 95%	
(Catto et al., 2003)	ANN	Bca	Survival predictor	Age, gender, grade, smoking, previous cancer, p53, hMLH1, hMSH2	Survival in months	109 patients from one centre with bladder TCC	10% cross validation RMS	RMS = 11.7	
(Parekattil et al., 2003)	ANN	Bca	diagnosis	Urine levels of nuclear matrix protein-22, monocyte chemoattractant protein-1 and urinary intercellular adhesion molecule-2	Cancer and benign	253 Data from one cystoscopy clinic	ROC, Sp, Se	Se 100%, Sp 75.7%, PPV 32.9%, NPV 100%,	The three factors improve diagnosis
(Djavan et al., 2004)	ANN	BPE	Significant LUT symptoms in BPE	Age, PSA, Qmax, TZV, TPV, Oss, ISS, PVR	Progression or no	397 patient with mild LUTS from 4 centres	1/3 CV ROC, Sp, Se, Then sensitivity analysis	Ac 79%, Se 82%, Sp 77%, PPV 78%, NPV 81%	PSA, Oss, TZV are correlated to disease progression
(Kshirsagar et al., 2006)	ANN	Hgon	Diagnosis of hypogonadism,	Age, ED, depression score, sexual health score, testosterone level	Risk of hypogonadism	148 one centre	70 test cases		Depression most significant, p<0.0019
(Stephan et al., 2005)	ANN	BPE/CaP	Diagnosis of BPE and CaP	Age, tPSA, %f PSA, TPV, MIC-1, Hk11, MIF	Cancer and benign	Single centre 371 patients	LOO	AUC 0.91, Se 90%, Sp 80%	Positive if all makers added together
(Abbod et al., 2006)	ANN	Bca	Survival and recurrence predictor	22 different genes variables	Risk and time to relapse	67 bladder neoplasms and 8 normal bladder specimens	Difference RMS 10 folds CV ROC AUC	RMS 5.2 Ac 100%	500 genes where reduced to 22 genes for creating the network, thus significant
(Abbod et al., 2006)	FNM	Bca	Survival and recurrence predictor	66 rules from 11 gene variables	Risk and time to relapse	67 bladder neoplasms and 8 normal bladder specimens	Difference RMS 10 folds CV ROC AUC	RMS 2.2 Ac 100%	500 genes where reduced to 22 genes for creating the network, thus significant
(Catto et	FNM	Bca	Recurrence (classifier)	Age, gender, grade, smoking, previous cancer, p53, methylation index (% of loci on	Recurrence or not	117 patients with 1ry TCC or UCC from one centre	10% cross validation ROC, LR	AUC 0.98, Se 88-100%, Sp 94-100%, Ac 100%	p value calculated to compare all models, the effect of combining HK p53 with other variables

al., 2006)				chromosomes), RARB methylation					
(Catto et al., 2006)	FNM	Bca	Survival predictor	Age, gender, grade, smoking, previous cancer, p53, methylation index (% of loci on chromosomes), RARB methylation	Survival in months	117 patients with 1ry TCC or UCC from one centre	10% CV Kaplan Maier for survival	Average error = 5 months	Interrogate different markers to suggest a predicative combination
(Catto et al., 2006)	ANN	Bca	Recurrence (classifier)	Age, gender, grade, smoking, previous cancer, p53, methylation index (% of loci on chromosomes), RARB methylation	Recurrence or not	117 patients with 1ry TCC or UCC from one centre	10% cross validation ROC, LR	Ac 89-90%, Se 81-87%, Sp 95-100%	
(Catto et al., 2006)	ANN	Bca	Survival predictor	Age, gender, grade, smoking, previous cancer, p53, methylation index (% of loci on chromosomes), RARB methylation	Survival in months	117 patients with 1ry TCC or UCC from one centre	10% CV Kaplan Maier for survival	Average error = 9 months	p for comparison ANN & FNM calculated
(Stephan et al., 2006 b)	ANN	CaP	Diagnosis of cancer in PSA 1-4 4-10	Age, tPSA, %fPSA, TPV, DRE, -5pro PSA, -7, pro PSA	Risk of cancer	2 centre PSA 1-10 & TRUS 6-12 cores, 898 patients	ROC, Spearman correlation co efficient LOO	AUC 84%	Pro PSA improved detection rate in 1-4 and improved %fPSA performance in 4-10 group
(Stephan et al., 2006 a)	ANN	CaP	Early CaP diagnosis	Age, tPSA, %fPSA, hK11, hK11/tPSA , hK11/%tPSA	Cancer or benign	357 with histologically proven cancer or BPE	ROC Se, Sp test set 206 with histologically proven cancer or BPE	AUC 0.84, Se 90%, Sp 52%	Sensitivity analysis of these variables to demonstrate their impact on AUC
(Stephan et al., 2008)	ANN	CaP	Early CaP diagnosis	Age, tPSA, %f PSA, TPV, DRE (PSA done by five different assays)	Risk of cancer	585 patients with suspected cancer PSA 0.49-27	ROC AUC 25% random set 195 patients and LOO	AUC 0.91 (mean value)	Authors suggests developing PSA assay specific ANN to optimise function
(Cinar et al., 2009)	ANN	CaP	Prostate cancer early diagnosis	Age, BMI, tPSA, fPSA, TPV, PSAD, smoking, systolic-diastolic pressure, pulse, Gl	Cancer or benign	300 patients' data with suspected cancer from one centre	10- folds CV ROC Se, Sp	Ac 79%, Se 81%, Sp 78%	
(Cinar et al., 2009)	SVM	CaP	Prostate cancer early diagnosis	Age, BMI, tPSA, fPSA, TPV, PSAD, smoking, systolic-diastolic pressure, pulse, Gl	Cancer or benign	300 patients' data with suspected cancer from one centre	10- folds CV ROC Se, Sp	Ac 81%, Se 84%, Sp 75%	Smoking is a significant classifier but not BMI

(Stephan et al., 2010)	ANN	CaP	Diagnosis	Age, tPSA, %f PSA, DRE, TPV	Risk of cancer	PSA2-20 393proscpective data	ROC AUC LOO	AUC 0.75, Se 90%, Sp 37%	Demonstrate the impact of different data cohorts on ANN performance
(Catto et al., 2010)	FNMANN	Bca	Gene micro array to predict UCC progression	200 genes reduced from 2800 by Pearson correlation	Cancer progression to muscle invasive or metastatic	66 tumours from 34 patients in one centre	COX multivariate analysis 10 folds CV	11 new gene signatures	200 gene micro array reduced to 11 gene signatures
(Serati et al., 2011)	ANN	UDyn	Urodynamic interpretation	Age, BMI, menopause, sexual activity, UTI, number of vaginal deliveries, surgery,	U Dyn diagnosis	802 data from single centre POP with symptoms and UDS performed	ROC and compare to multi linear regression CV 20%	AUC 80% (Average)	ANN cannot replace Urodynamic
(Gil et al., 2012)	ANN	Fert	Seminal profile from questionnaire about life habits and health status	Age, BMI, demographic, medical history facts, smoking, alcohol, life style and cloth.	Seminal profile	100 volunteers one centre study	ROC AUC Se, 10 Folds cross validation	Se 73-94%, Sp 25-45%, PPV 79-92%, NPV 7.4-54%	Comparison of different AI classifiers with same variables
(Gil et al., 2012)	SVM	Fert	Seminal profile from questionnaire about life habits and health status	Age, BMI, demographic, medical history facts, smoking, alcohol, life style and cloth.	Seminal profile	100 volunteers one centre study	ROC AUC Se, 10-fold CV	Se 74_99%, Sp 12-21%, PPV 75-91%, NPV 4-86%	
(Gil et al., 2012)	DT	Fert	Seminal profile from questionnaire about life habits and health status	Age, BMI, demographic, medical history facts, smoking, alcohol, life style and cloth.	Seminal profile	100 volunteers one centre study	ROC AUC Se, 10-fold CV	Se 72-96%, Sp 12-41%, PPV 77-90%, NPV 4-48%	
(Girela et al., 2013)	ANN	Fert	Seminal profile from questionnaire about life habits and health status	Age, season, childhood disease, surgery, trauma, smoking, alcohol, hours sitting ANNA1	Sperm concentration	100 volunteers one centre study	ROC AUC Se, Sp 10 Folds CV	Se 95%, Sp 50%, PPV 93%, NPV 60%	
(Girela et al., 2013)	ANN	Fert	Seminal profile from questionnaire about life habits and health status	Age, BMI, marital status, vaccines, siblings, allergy, baths, hours of sleep ANNA2	Sperm motility	100 volunteers one centre study	ROC AUC Se, Sp	Se 89%, Sp 44%, PPV 89%, NPV 44%	

(Stephan et al., 2013)	ANN	CaP	Statistical evaluation of PSA INDEX	Age, TPV, DRE, tPSA, %fPSA	Risk of Cancer	1362 from multiple centres with suspected CaP and PSA 1.6-8.0	ROC AUC and comparison to other markers	AUC 0.7 - 0.74	
(Stephan et al., 2013)	ANN	CaP	Statistical evaluation of PSA INDEX	Age, TPV, DRE, tPSA, %fPSA, %p2PSA	Risk of Cancer	1362 from multiple centres with suspected CaP and PSA 1.6-8.0	ROC AUC and comparison to other markers	AUC 0.73 - 0.79	
(Stephan et al., 2013)	ANN	CaP	Statistical evaluation of PSA INDEX	Age, TPV, DRE, tPSA, %fPSA, %fPSA prostate health index (p2PSA / fPSA X square root tPSA)	Risk of Cancer	1362 from multiple centres with suspected CaP and PSA 1.6-8.0	ROC AUC and comparison to other markers	AUC 0.73- 0.8	Prostate Health index improved ANN performance
(Wang et al., 2015)	ANN	Bca	Survival post cystectomy	Age, gender, albumin, surgical approach, tumour stage, follow up period, type of diversion	5 years survival	117 patients with post cystectomy from one centre	ROC, Se, Sp Ac, 10 Folds cross validation	Ac 72-80%	Comparison of 7 different machine learning. RELM & ELM had best performance
(Cai et al., 2011)	ANN	CaP	+ve lymph nodes to the total number of lymph nodes in predicting BCF	Age, tPSA, Clinical stage, Gl, seminal vesicle invasion, number of positive lymph nodes and laterality of lymph node involvement.	BCF	124 cases with lymph node dissection	hazard ration for each variable	LND, Gl, and stage were identified as independent prognostic	LND is more prognostic than their number
(Kim et al., 2014)	BN	BPE	Correlation between symptoms, decision and outcome of surgery	Age, Qmax, PVR, PSA, TPV, TZV, BOO on UDS, and IPSS scores (stratified)	surgical decision-BN model, the outcome of surgery	1108 cases from one centre	ROC AUC and correlation coefficient	AUC 0.8 TZV (R=0.396, P<0.001), treating physician (R=0.340, P<0.001) and BOO on UDS (R=0.300, P<0.001)	TPV, physician, BOO on UDS, and the IPSS item of intermittency were factors that directly influenced Decision-making in physicians treating patients with LUTS/BPE.
(Green et al., 2016)	ANN	CaP	Progression biomarkers	Gene microarray	Cancer progression and DSS	192 tissue histology results	MSE for each variable, then Kaplan Meyer and Pearson's x2-tests	10 gene microarrays identified by ANN	Ki67 and DLX2, appear to predict CaP specific survival and metastasis.
(Logvinenko et al., 2015)	ANN	VUR	Renal ultrasound to predict voiding cystourethrogram (VCUG)	Renal ultrasound findings	abnormal VCUG	2259 cases post UTI and had VCUG	ROC AUC	Se 64.2%, Sp 59.6%, PPV 61.6%, NPV 62.2%, AUC 0.6852	Renal ultrasound is a poor screening test for VCUG-identified abnormalities.

.

2.3.1.6.1 Domains

Prostate cancer was once more the commonest domain in this application with a total of 15 systems analysing predictive factors for diagnosis of cancer, response to treatment and quality of life with prostatic disease.

One of the hot topics in Urological cancer is discovering alternative CaP diagnostic markers since serum PSA is not sensitive for distinguishing benign from malignant disease. *Stephan et al.* investigated the diagnostic value of three markers in this domain: Macrophage inhibitory cytokine-1, macrophage inhibitory factor and human kallikrein 11(Stephan et al., 2005). These were used as variables (nodes) in ANN models and compared their accuracy to the linear regression of %fPSA. They have reported that only the ANN model including all three variables was more accurate (AUC 91%, Se 90%, Sp 80%) than all other models proving his hypothesis that they are only relevant as when combined.

Similarly, another study estimated the predictive values of serum PSA precursors (-5, -7 proPSA) in diagnosing prostate cancer using and comparing the accuracy to %fPSA (Stephan et al., 2006b). The -5, -7 pro PSA were only significant in the cohort with PSA between 4-10 µg/l and did not improve the predictive accuracy when added to the %fPSA. The same author tested this hypothesis on another free PSA precursor (-2 proPSA) by developing ANN with the %p2PSA (-2 ProPSA: fPSA) among other disease variables, which have remarkably improved the system accuracy (AUC 85% from 75%)(Stephan et al., 2009).

Three systems evaluated the presence of bcl-2 and p53 (tumor suppressor genes) as a predictive variable for response to prostate cancer treatment (Naguib et al., 1998, Seker

et al., 2003). Their combination was reported to be significant (Ac 85%, $p<0.00001$) in Naguib et al. (1998) but Seker et al. (2003) found that only bcl-2 is relevant in the other two models (accuracy 63-68%).

Bladder cancer diagnosis and disease progression was the second most common domain with 13 systems. Parekattil et al. (2003) have modeled an ANN with three novel urine markers: urine levels of nuclear matrix protein-22, monocyte chemoattractant protein-1 and urinary intercellular adhesion molecule-1, to predict the diagnosis of bladder cancer and it succeeded in predicting all cancer free patients when the three variables were used as a group. Catto.et al. Catto et al. (2010) developed two AI models (ANN & FNM) performing microarray analysis on genes associated with bladder cancer progression. Their models narrowed down these genes from 200 to 11 progression-associated genes out of 200 ([OR]: 0.70; 95% [CI], 0.56–0.87), which were found to be more accurate than the regression analysis when compared to the specimen immunohistology results.

Parekattil et al. (2003) model predicting the prehistology diagnosis of malignancy based on urine level of novel tumour markers. Their ANN was found to be more accurate (Se 100%, Sp 75.7%) than haematuria diagnosed on urine dipstick (Se 92.6%, Sp 51.8%) and atypical urine cytology (Se 66.7%, Sp 81%).

ESWL of renal stones was the research domain of Poulakis et al. (2002) Michaels et al. (1998), where they aimed at identifying significant variables correlated to the treatment outcome (stone free) and developing a predictive model. Michaels et al. (1998) model did not recognise residual fragments following ESWL as a significant risk for triggering further stone growth and Poulakis et al. (2002) identified these factor:

positive BMI, infundibular width (IW) 5mm, infundibular ureteropelvic angle 45% or more (IUPA), to be all predictive of lower pole stone breaking and clearance.

Benign prostatic hyperplasia was modelled in a system (Djavan et al., 2004) to link the disease specific clinical and radiological factors with the disease progression in patients with mild disease (IPSS <7) and not receiving any treatment. His ANN identified: obstructive symptoms (Oss), PSA of more than 1.5 ng/ml and transitional zone volume of more than 25 cm^3, to be correlated to disease progression and can accurately predict 78% of the cohort who will need further treatment.

Urinary dysfunction diagnosis accuracy by clinical symptoms was compared to urodynamic findings in female patients with pelvic organ prolapse by Serati et al. (2011) and both the linear regression and ANN models could not establish relation between the symptoms and urodynamic based diagnosis hence dismissing the hypothesis of only relying on clinical symptoms to reach an accurate diagnosis and replace the need for urodynamics study.

Hypogonadism (Hgon) was represented in Kshirsagar et al. (2006) system where the diagnosis was made based on patient's age, erectile dysfunction and depression with AUC of 70% ($p<0.01$).

2.3.1.7 Image analysis

This one of the advancing applications of AI in medicine where the system either analyse the variables in the reported medical images as data input or identifies these variables through a separate image analyser without the need for expert to report the scan or images. The first category was included among other systems mentioned above as in the diagnosis prediction domain where Ronco and Fernandez (1999) included different variables from TRUS in the system input to predict CaP diagnosis. In this

domain, we focused on the other group where the images are presented to the machine in the form raw data translated by the image analyser and the system will then apply their machine learning to identify the cause effect pattern.

Table 2.9 Image analysis

Expert Systems in this application analysed images from histology and radiological scans to learn patterns that are correlated to a specific diagnosis. They have proven to be effective in this domain and they facilitated diagnosis of cancer and even delivering radiotherapy dosage.

Art	Mdl	Dom	Subdomain	Variables	Output	System training	Validation	Statistical outcome
(Wells and Niederer, 1998)	ANN	CaP	Radiotherapy dose planning	Patient prostate contour points (anterior, posterior and 5 lateral)	Anterior, posterior and lateral beam	12-68 patients record of radiotherapy treatment planning	Average asymmetry of ANN and acceptance by dosimetrists Small field prostate (n=133) and for large field prostate (n=64)	Average asymmetry of ANN 0.20% and acceptance by dosimetrists was 96% (small field prostate) and 88% for large field prostate
(Loch et al., 1999)	ANN	CaP	Diagnosis of localised disease from TRUS	Pixel distribution and grey levels of the TRUS images	Benign, malignant with Gleason grading	53 images of benign and malignant sample images from 5 patients	Compare to histology results of 500 pictures from 61 patients post RRP for localised disease in one centre	Sp 99%, Se 83%, true positive for isoechoic is 97%
(Mattfeldt et al., 1999)	ANN LDA	CaP	progression post RPP	Prostate volume, PSA, Pathology morphometric variables LDA	Progression or no	Progression t2n0 post RRP, 228 patients from one centre	ROC, Sp, Se, LOO 39 cases	Ac 70%, Se 55%, Sp 85%
(Mattfeldt et al., 1999)	ANN LVQ	CaP	progression post RPP	Prostate volume, PSA, Pathology morphometric variables	Progression or no	Progression t2n0 post RRP, 228 patients from one centre	ROC, Sp, Se, LOO 39 cases	Ac 90%, Se 95%, Sp 85%,
(Mattfeldt et al., 1999)	ANN LVQ PAK	CaP	progression post RPP	Prostate volume, PSA, Pathology morphometric variables	Progression or no	Progression t2n0 post RRP, 228 patients from one centre	ROC, Sp, Se, LOO 39 cases	Ac 83%, Se 85%, Sp 80%
(Mattfeldt et al., 1999)	ANN MLF-bp	CaP	progression post RPP	Prostate volume, PSA, Pathology morphometric variables	Progression or no	Progression t2n0 post RRP, 228 patients from one centre	ROC, Sp, Se, LOO 39 cases	Ac 76%, Se 73%, Sp 80%
(Llobet et al., 2007)	kNN	CaP	TRUS cancer image analysis	Image pixels segmented by tissue descriptor (spatial grey level dependence)	Predict cancer	Images of 202 patients with suspected CaP at one centre	87 randomly selected patients Comparison to other classifiers and ROC	AUC 0.6

(Hassanien et al., 2011)	ANN	CaP	TRUS Image segmentation	Pixel's colour values from TRUS images	TRUS image segmentation	212 CaP TRUS data	Overlap measure (compared to expert segmented boundary) on 10 random images	81% mean overlap measurement
(Matulewicz et al., 2014)	ANN	CaP	MRI cancer diagnosis	256 MRSI spectra (resonance intensities at given PPM)	Cancer or benign	5308 voxels of 18 patients with CaP in a retrospective study	15% validation ROC Se, Sp	AUC 0.95, Se 50%, Sp 99%
(Matulewicz et al., 2014)	ANN	CaP	MRI cancer diagnosis	256 MRSI spectra (resonance intensities at given PPM), peripheral & transition zone, periurethral & outside region	Cancer or benign	5308 voxels of 18 patients with CaP in a retrospective study	15% CV validation ROC Se, Sp	AUC 0.97, Se 62%, Sp 99%
(Gatidis et al., 2015)	SVM	CaP	Diagnosis of cancer from pMRI images	Image segmentation then clustering voxels	Cancer or benign	16 pMRI images with CaP	Correlation coefficients of voxel parameters	Mean accuracy of 84%
(Pantazopoulos et al., 1998b)	ANN	Bca	Image histology analysis	Image histology analysis (measurements of the segmentation of nuclei and other features)	Benign and malignant	141 randomly chosen cell images (30%)	329 cell images (70%) ROC, Sp, Se	Sp 100%, Se 82%, PPV 96%, NPV 80%, Ac 88%
(Xiao et al., 2016)	FCM	Bca	Diagnosis of tumour	Bladder wall segmentation and tumour region extraction, To detect bladder abnormalities, four volume-based morphological features: bent rate, shape index, wall thickness, and bent rate difference between the inner and outer surfaces	Bladder neoplasm	16 Bladder tumour MRI images	Overlap Ratio (OR)	OR 86.3 %
(Hurst et al., 1997)	ANN	Bca	Transitional cell cytology analysis	Cytology image analysis and pixel variations as variables	Cancer or benign	16 cytology images	comparison to experts, x2 test	75% concordance with the experts
(Volmer et al., 1994)	ANN	Nlt	Spectroscopy stone analysis	Absorbance infra-red spectrum of 91 wave lengths	Stone composition	160 and 57 stone mixtures	Predictive accuracy, root mean square error on 36 independent stone samples	Overall good predictive value

2.3.1.7.1 Domains

Prostate cancer image analysis was modelled in 10 systems to enhance diagnostic accuracy as in Hassanien et al. (2011) and disease progression prediction as in Mattfeldt et al. (1999). The first system represented each TRUS image pixel as one variable or neuron in a pulse coupled neural network and trained their system with 212 prostate cancer images to segment prostate gland boundary with an average overlap accuracy (overlap measure = difference between PCNN boundary and the expert) of 81% for ten images (Hassanien et al., 2011).

The other 4 systems analysed histological images of a cohort of patients post RP with clinically localised CaP to predict the disease progression. The histological images were given coloured coding and analysed by the system that used variables as % of epithelial cell and glandular Lumina to identify the high risk group for disease recurrence with an accuracy reaching 90% (Mattfeldt et al., 1999).

LUT disease urine cytology images were analysed by 2 models in Pantazopoulos et al. (1998a), which identified all patients with benign disease with an overall accuracy of 97%.

Nephrolithiasis stone biochemistry analysis can be achieved through an expert analysis of infrared spectroscopy which was simulated by Volmer et al. (1994) where the infrared spectra wavelength numbers were modelled as input variables and the system prediction accuracy of the expert analysed stone specimen had a root square mean error of 3.471.

2.3.2 **Qualitative analysis**

The same articles were considered for the qualitative analysis against the four stages (validation, verification, evaluations and credibility) reported in Okeefe industrial

survey (Okeefe and Oleary, 1993) and Benbasat article (Benbasat and Dhaliwal, 1989). The completion of the four stages examined in this qualitative analysis was demonstrated by none of the included systems. There is a possibility that some of these missing stages has been performed but not published in the scientific literature.

Validation was performed by almost all of the systems (166 out 169) with varying degree of study strength, bias and limitations (Table 2.10). Most of the data driven systems (ANN, SVM, BN, kNN and FNM) were validated by the ROC and AUC by having a training and validation set or cross validation or applying the leave one out technique. Samli et al. enhanced the validity of their system by estimating the kappa statistics with the ROC (Samli and Dogan, 2004).

Evaluation was only performed by a small fraction of these systems (n=6). Their evaluation was aiming at the user or the expert but rarely both. There is no evidence to support that these were performed at early stages to determine the substantiality of the system to the user.

System credibility and verification were never performed. It would be implied that the verification was performed to an extent but not reported as it is a technical part of the development.

System development limitation and bias evaluation demonstrated a majority of acceptable validation methodologies with valid statistical analysis. However, a number of observed limitations (Table 2.10) were reported with the common encounter being the consideration human opinion as a gold standard (n=9). For instance, the gold standard in diagnosing prostate cancer is tissue biopsy confirmation. The interpretation of the expert clinical diagnosis as the gold standard reference can lead to statistical errors and invalidate the study.

Table 2.10 Qualitative assessment of urological expert systems

All systems' development was qualitatively assessed against the common industrial steps in the development pathway described by Okeefe and Benbasat. With exception of the system validation, the rest of the cycle was defective with no explanation. The validation had variable degree of strength with common application of the receiver operator characteristic for estimating the area under the curve for data driven systems.

Art	Mdl	Validation methods	Credibility	evaluation	validation	verification	Strength and bias
(Binik et al., 1988)	RBR	patients' evaluation	No	Yes	Yes	No	Only qualitative evaluation
(Petrucci et al., 1991)	RBR	blinded comparison against 4 experts with independent experts rating and 3 centres RCT pilot trial	Yes	Yes	Yes	No	Consideration of system evaluation with real time testing but small number
(Sucevic and Ilic, 1991)	FRB	improve practitioner accuracy	No	No	No	No	Insufficient info on development and validation
(Gorman, 1995)	RBR	RCT reliability and validity by experts' reviews	Yes	Yes	Yes	No	Small number in the study and short duration of follow up
(Bagli et al., 1998)	ANN	ROC, Sp, Se	No	No	Yes	No	Small number for validation
(Chang et al., 1999)	FSS	ROC, Sp, Se	No	No	Yes	No	2 methods for validation, compared to experts and data
(Loch et al., 1999)	ANN	Compare to histology results	No	No	Yes	No	No comparison to human to demonstrate usability, no p value or CI
(Catto et al., 2003)	FNM	ROC, LR, RMS	No	No	Yes	No	p value calculated to compare all models
(Catto et al., 2003)	ANN	ROC, LR, RMS	No	No	Yes	No	p value calculated to compare all models, the effect of combining HK p53 with other variables
(Fujikawa et al., 2003)	ANN	ROC, Sp, Se	No	No	Yes	No	No p value
(Hamid et al., 2003)	ANN	Correlation co-efficient	No	No	Yes	No	Correlation co-efficient between expert and system? Kappa more accurate
(Saritas et al., 2003)	FRB	Not published	No	No	No	No	Not validated
(Zlotta et al., 2003)	ANN	AUC ROC	No	No	Yes	No	p value calculated vs LR
(Boyington et al., 2004)	RBR	feedback from patients with no control group	No	Yes	No	No	No validation but user (patient evaluation)

(Koutsojannis and Hatzilygeroudis, 2004)	FRB	comparison to experts and non-experts	No	No	Yes	No	Expert as gold standard
(Pereira et al., 2004)	RBR	PPV 62%, NPV 100% Se 100% Sp 33%	No	No	Yes	No	Small number, low specificity
(Samli and Dogan, 2004)	ANN	ROC AUC then compare with LR, kappa stats	No	No	Yes	No	Multimodal of validation
(Serrano-Durba et al., 2004)	ANN	ROC, Sp, Se	No	No	Yes	No	Not long term follows up
(Lee et al., 2006)	ANN	ROC (0.74 & 0.86)	No	No	Yes	No	TRUS finding from expert panel, human as gold standard
(Catto et al., 2006)	FNM	ROC, LR	No	No	Yes	No	p value calculated to compare all models
(Catto et al., 2006)	ANN	Kaplan Maier for survival	No	No	Yes	No	p for comparison ANN & FNM calculated
(Llobet et al., 2007)	kNN	Comparison to other classifiers and ROC	No	Yes	Yes	No	Evaluated the usability of the product and was found to have less than significant effect.
(Stephan et al., 2006a)	ANN	ROC Se, Sp	No	No	Yes	No	Sensitivity analysis of input variables
(Altunay et al., 2009)	ANN	ROC 0.7, accuracy 79%	No	No	Yes	No	Compare to experts without accounting for human error
(de Paula Castanho et al., 2008)	FRB	ROC Se, Sp	No	No	Yes	No	No user evaluation
(Koutsojannis et al., 2008)	FRB	Ac 0.76, Se 0.79, Sp 0.75	No	No	Yes	No	Expert as gold standard
(El-Mekresh et al., 2009)	ANN	ROC Compare to LR	No	No	Yes	No	CI calculated
(Koutsojannis et al., 2009)	FRB	Ac 0.93, Se 0.97, Sp 0.99	No	No	Yes	No	Expert as gold standard
(Kolasa et al., 2009)	ANN	Prediction error percent	No	No	Yes	No	Experimental results
(Lee et al., 2010)	SVM	ROC AUC	No	No	Yes	No	P value calculated to compare all models
(Hassanien et al., 2011)	ANN	Overlap measure (segmented by experts)	No	No	Yes	No	Expert as gold standard
(Gil and Johnsson, 2010)	ANN	Ac 0.84, Se 0.93, Sp 0.33	No	No	Yes	No	Experts verified data no account for human error
(Papageorgiou, 2012)	FNM	Accuracy 86.8 %	No	No	Yes	No	Guidelines as gold standard

(Koutsojannis et al., 2012)	RBR	Evaluation by experts, 95 retrospective	No	No	Yes	No	Expert as gold standard, qualitative evaluation
(Torshizi et al., 2014)	HYB FUZ ZY ONT	kappa vs experts, k=0.89	No	No	Yes	No	Kappa limitation prospective, randomisation,
(Hao et al., 2013)	RBR	Se 0.95, Sp 0.72, Bayesian analysis S&S, usability of system by Likert scale (Cronbach's alpha 0.9)	Yes	Yes	Yes	No	Full system evaluation but nurse as gold standard, no attempts to eliminate error
(Tsao et al., 2014)	ANN	ROC AUC compare with Partin nomogram and LR	No	No	Yes	No	Not correlation with user
(Lopes et al., 2013)	FNM	kappa vs experts, Se 0.95, Sp 0.92	No	No	Yes	No	Human expert as gold standard and no qualitative evaluation (weight of error)
(Wadie et al., 2006)	ANN	Ac 60% (testing) 75% (training)	No	No	Yes	No	Compare to gold standard, Urodynamic
(Tanthanuch and Tanthanuch, 2004)	ANN	PPV 100%	No	No	Yes	No	No calculation of NPP and overall accuracy
(Goyal et al., 2007)	FNM	Correlation coefficient = 0.99	No	No	Yes	No	Small number of cases for validation
(Xiao et al., 2016)	FCM	OR 86.3 %	No	No	Yes	No	Comparison with experts as gold standard than mapping to histology
(Logvinenko et al., 2015)	ANN	ROC, Se 64.2%, Sp 59.6%, PPV 61.6%, NPV 62.2%, AUC 0.6852	No	No	Yes	No	Similar to urodynamic as research tool
(Yuksel et al., 2013)	FRB	None	No	No	Yes	No	No validation

2.4 **Discussion**

There are many ES available in Urological domains, with a wide range of models, applications, domains and target users to include patients, students, non-experts, experts and researchers. The number of published systems has risen over the years but with a consistent lack of publications to reporting their real time testing or implementation in the healthcare. Previous review articles aimed to familiarise the readers with ES existence and application (Lamb and Niederberger, 1993, Kattan et al., 1996) while the majority focused on the evolution and application of one ES models (artificial neural network) in a defined domain predominantly prostate cancer diagnosis. However, their scope did not include the ES uptake gap and mismatch between production and implementation. This study quantified their existence, application while examining their life cycle.

CaP was evidently the commonest domain in almost all applications contributing with more than two thirds of the systems (91 systems in total). Different aspects of this domain have been simulated by these systems to include diagnosis, therapeutics, predictions of disease progression or treatment outcome, researching variables and medical images analysis. Most of these systems were simulating urologist cognitive function with little guidance on their benefits and how they can be implemented to improve cancer decision making. In industry, this is usually performed before the system development by evaluating the system usability from the user perspective. This part has lacked in all of those systems and are possibly a core reason for the lack of their integration in urological health care.

Furthermore, none of these systems has been a subject to live testing in a well-designed study to prove its efficacy over standard tools or in the clinical context to prove its

validity and utility bearing in mind their complex structure and function. Some of the existing tools are less complex and in a head to head comparison of modern regression models and the ANN, they were not found to be superior Kawakami et al. (2008).

There was a wide diversity of modelling in Urological ES with ANN being the most common model. They bypass the need for direct learning from experts and the exhaustive process of knowledge acquisition which is a core requirement for knowledge based systems, and can arrest the whole system progress (Pandey and Mishra, 2009). However, their analytical hidden layer of nodes "black box phenomenon" has been a subject for wide criticism and rejection from clinicians due to lack of transparency and understanding of its function. Stephan *et al.* suggested a statistical solution to identify the variables significance by performing sensitivity analysis (Stephan et al., 2006a). This has been reported by estimating the variation of the AUC with introduction or elimination of each variable. This can only reflect the significance of each variable but does not explain how the cases are being solved. This can be useful in research as they can identify significant variables in noise contaminated data. This can be effectively applied in the field of academic urology (Catto et al., 2010) where the system successfully identified the relevant gene signature for bladder cancer progression which saved time and cost of gene microarray analysis of all suspected genes.

The qualitative analysis demonstrated that validation is the only stage of the development cycle to be applied by most of the systems and there is a lack of system evaluation, credibility and verification. The evaluation can be subdivided into usability (usually by average user), utility and system quality (by experts) (Benbasat and Dhaliwal, 1989). Despite being a crucial stage of ES development, there has been a lack of attention in the published articles to integrate it into the development life cycle. This

can mean the whole system can fail and also challenge its uptake (Okeefe and Oleary, 1993)..

The application of the utility evaluation on the prostate cancer diagnosis as a domain would demonstrate the lack of community-based systems to refer patients with suspected cancer then manage patients with stable disease, where NICE have identified a need for such a system (NICE, 2014, NICE, 2008).

Another common aspect lacking in these articles was the coupling of their system development methodology with the medical device registration requirements and lack of attention to this can limit their ES introduction into the health care (Ammenwerth et al., 2013).

2.5 Conclusion

ES have been advancing in Urology with demonstrated versatility and efficacy. They have suffered from lack of formality in their development, testing and methodology for registration, which has limited their uptake. Future research is recommended in identifying criteria for successful functional domain applications, knowledge engineering and integrating the system development with the registration requirement for their future implementation in the health care systems.

Chapter three - Prostate cancer knowledge acquisition (declarative knowledge)

3.1 Synopsis

3.1.1 Background

Prostate cancer has a rising incidence and significant impact on health care. The disease is usually diagnosed and treated in secondary and tertiary centres by disease control methods (watchful waiting, hormonal ablation therapy, focal therapy, chemotherapy) or disease cure (radical radiotherapy, radical surgical prostatectomy) which can follow a period of active surveillance. Following disease stabilisation, the patient receives their follow up care in either the primary or secondary care which have wide regional variance.

3.1.2 Objectives

1. To demonstrate the impact of prostate cancer on health care and review evidence from the literature and guidelines for treatment strategies and follow up.
2. Elicit relevant of declarative (factual) knowledge required for building the knowledge base of the follow up domain.
3. Identify current available follow up service for stable prostate cancer and the potential impact of computer led follow up on patients' care.

3.1.3 Results

Prostate cancer care has an increasing impact on health care. Current guidelines have consensus on definitions of biochemical failure post radical treatment and available treatment for disease recurrence. Treatment with chemotherapy is evolving for non-curative and palliative disease.

The follow up of the stable disease should be delivered in the community but this does not happen as much as it should do, which could reflect the lack of the required expertise or resource.

3.2 Introduction

Prostate cancer is the most common cancer to be diagnosed in men with an incident rate (European age standardised) of 182 per 100,000 in 2013 (CancerResearchUK, 2016a). There has been a steady increased incidence in both Europe and the UK, (Figure 3.1) which is believed to be a consequence of the increase uptake of PSA checking in the 1980s, and the incidental diagnosis after transurethral resection of the prostate (TURP) (Bray et al., 2010, Quinn and Babb, 2002, Brewster et al., 2000). In fact, it has been predicted that this incidence will demonstrate a 7 fold increase in figures by 2030 (Mistry et al., 2011).

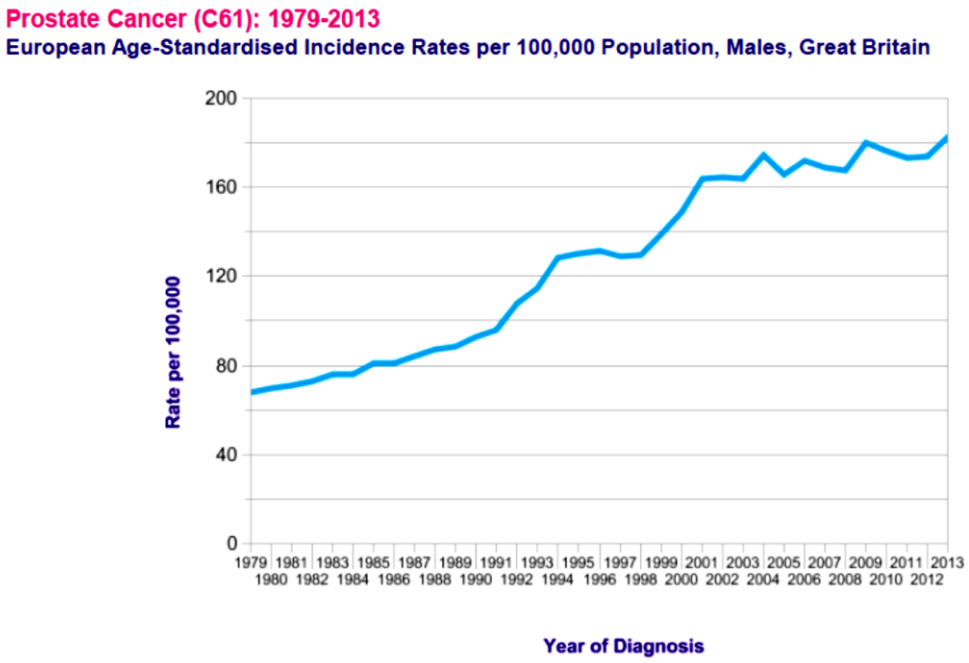


Figure 3.1 European Age Standardised incidence rate of prostate cancer in UK years
Graph copied from Cancer Research UK demonstrating the steady rise in incidence of prostate cancer which is attributed to increase PSA checking and incidental finding post TURP.

3.3 Risk factors

Prostate cancer incidence increases significantly with age. There is a 5-folds increased incidence between age of 55 and 79 followed by a slight decline (Figure 3.2)

(CancerResearchUK, 2016a) .This is may be a reflection of the decline of PSA checking and prostate biopsies in this age group (Williams et al., 2011).

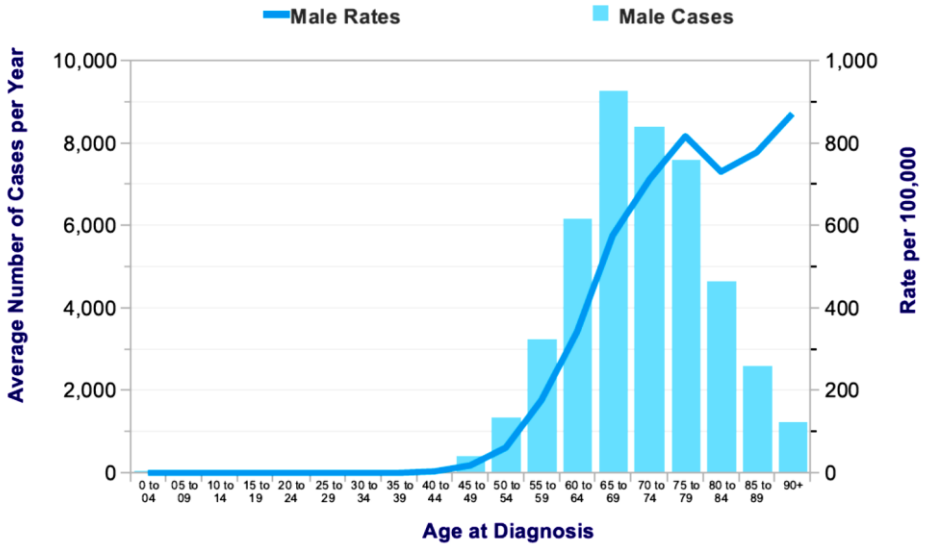


Figure 3.2 Prostate Cancer, Average Number of New Cases per Year & Age-Specific Incidence Rates, UK
Age has strong correlation with age and the decline at the age of 80 is attributed to lack of PSA checking and prostate biopsies.

Family history of prostate cancer is strongly correlated with increased incidence rate of the disease specially with first-degree relatives (Bruner et al., 2003, Johns and Houlston, 2003, Kicinski et al., 2011). The brother- brother relative risk was found to be more significant (2.87, 3.14 & 3.4) than the father-son (2.14, 2.35 & 2.5) (Kicinski et al., 2011, Johns and Houlston, 2003, Bruner et al., 2003). The risk is even higher if the relative was diagnosed before the age of 65 or when two or more first-degree relatives are affected (Johns and Houlston, 2003, Kicinski et al., 2011, Hemminki and Chen, 2005). The disease tends to be more common in the black race when compared to the white but the later was found to have a higher incidence than the Asians and Chinese (NCIN, 2009).

DNA damage repair genes mutations are correlated with prostate cancer specially the BRCA 2 mutations (Breast Canc Linkage, 1999, Bancroft et al., 2014). The BRCA 2 carriers tend to have more aggressive disease and worse outcome (Mitra et al., 2008).

Other known risk factors linked to increased risk are height (Zuccolo et al., 2008), previous renal cell cancer (Liu et al., 2011, Neuzillet et al., 2007) and insulin like growth factor1 (Renehan et al., 2004, Roddam et al., 2008, Rowlands et al., 2009, Price et al., 2012)

3.4 **Prostate cancer mortality**

Prostate cancer is the second common cause of cancer related death among the male population in the UK, with an age specific mortality rate of 48.1 per 100,000 in 2014 (CancerResearchUK, 2016b). The mortality rates are age related with an upward trend starting from the late 50s to reach the highest figure by the age of 85+ (CancerResearchUK, 2016b) Figure 3.3.

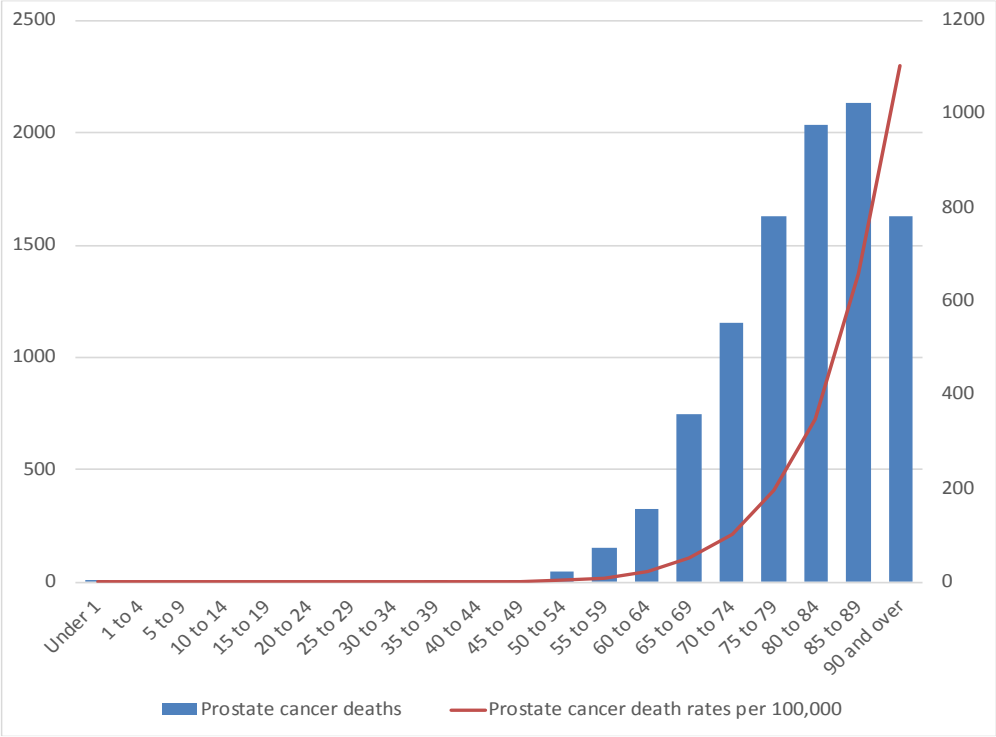


Figure 3.3 Number of deaths from cancer and age-specific rates of deaths from cancer (per 100,000) in England, 2016. Positive correlation of prostate cancer disease specific mortality and age.
Source: National Cancer Registration and Analysis Service within Public Health England; Office for National Statistics

The trend of mortality has changed over the last few decades showing an initial rise between the 70s and the early 90s then consistent decline (CancerResearchUK, 2016b). There is a strong belief that this pattern is due to early diagnosis from increase of PSA testing and the improved treatment strategies (Collin et al., 2008, Hanchanale et al., 2010).

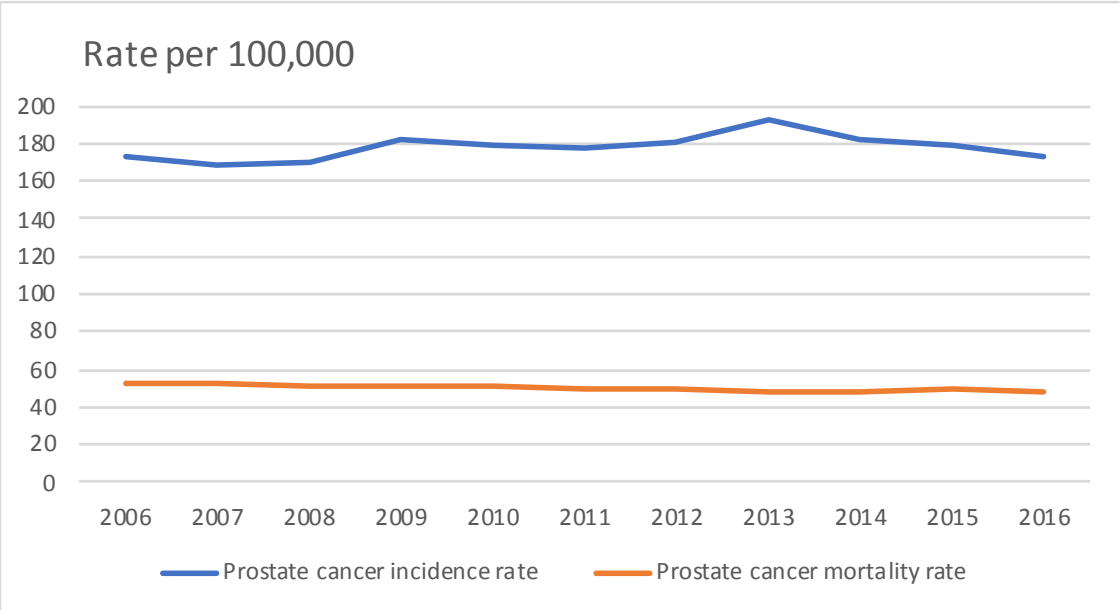


Figure 3.4 Directly age-standardised rates per 100,000 people of newly diagnosed cases of prostate cancer and cancer related mortality, England, 2006 to 2016, Source: National Cancer Registration and Analysis Service within Public Health England; Office for National Statistics
Mortality rates demonstrated evident improvement due to advances in disease diagnosis and treatment.(Collin et al., 2008, Hanchanale et al., 2010).

3.5 Treatment pathways

The National Institute of Clinical Excellence (NICE) 2019 guidelines categorised the newly diagnosed patients with prostate cancer into three groups: localised (organ confined), locally advanced and metastatic disease. The localised disease is further stratified into low, intermediate and high risk, according to PSA value, Gleason score and T score (Table 3.1). Following disease staging, the multidisciplinary team meeting recommends treatment options which are further discussed and tailored to the patient's

fitness and needs. The agreed treatment usually falls in one of five categories, watchful waiting, hormone ablation, active surveillance, radiotherapy and radical prostatectomy.

Table 3.1 Prostate cancer risk stratification, copied from NICE guidelines

Risk Group	PSA value at presentation	Gleason Score	Clinical stage
Low-risk	< 10 ng/ml	≤ 6	T1-T2a
Intermediate-risk	10–20 ng/ml	7	T2b
High-risk	> 20 ng/ml	8-10	T2c

3.5.1 **Watchful waiting**

In this pathway, the patients are not started on active treatment and are closely monitored for symptoms of disease progression. It is usually reserved for patients who are not candidate for curative radical treatment due to their age or co-morbidities and without symptoms of disease progression (NICE, 2014). When symptomatic, they can be started on other disease control treatment as hormonal treatment, palliative radiotherapy or other disease controlling models but not for disease cure strategies.

3.5.2 **Active surveillance**

The aim is to follow the progression of the low risk disease in patients who are candidate for radical treatment to avoid unnecessary radical interventions for indolent tumours. This is usually recommended for low risk localised disease (NICE, 2008), however the most recent NICE guidelines included offering AS to the intermediate risk group, for those whom have chosen deferred radical treatment (NICE, 2019). This has expanded the cohort suitable for AS and it was estimated that almost a half of the screened population with localised CaP can be suitable for deferred treatment (Albertsen, 2015). This contradicts an earlier study where the analysis of on British Association of Urological Surgeons (BAUS) cancer registry data identified only 9% of the patients to have low risk disease, suitable for AS (McVey et al., 2010). This was before the inclusion of the intermediate risk group and introduction of more refined diagnostic

tools as the multiparametric MRI (MP MRI) and transperineal template-guided biopsy of the prostate (TPTP) which have improved the accuracy of disease staging.

MP MRI has an increasing role to play in CaP diagnosis and can be used in different disease diagnostic stages. There is high level of evidence to support its use and it is recommended in suspected cancer (before the first biopsy), before the confirmatory 12 months biopsy and in AS follow up (Briganti et al., 2018). It has high value in identifying clinically significant and a large prospective multi-centre trial (PRECISION) reported the superiority of MRI targeted biopsy in both identifying clinically significant cancer and avoiding insignificant cancer (Kasivisvanathan et al., 2018). The MRI targeted biopsies can be done by image fusion software or just expertise of the user by mapping the MRI suspicious lesions to TRUS then target those with the biopsy needle. The targeted biopsy in the PRECISION study had a mean number of 4 core biopsies in contrast to the standard 12 cores biopsy which targets the periphery of the base, mid gland and apex and the former was superior in detecting clinically significant disease (adjusted difference, 12 percentage points; 95% confidence interval [CI], 4 to 20; P=0.005).

Even after a the first biopsy, MP MRI is still recommended specially if the biopsy confirmed localised low risk cancer as it can rule out any significant disease and can lead to upgrade of staging (Ahmed et al., 2017). There is no long term evidence to clarify the role on MP MRI in AS follow up and it should not substitute biopsies (Briganti et al., 2018). In fact the whole surveillance follow up protocol is nonstandard and vary according to the expertise recommendations due to the lack of compelling evidence. NICE contacted 210 health care professional and patients using a modified Delphi formal process and they reached a consensus on AS protocol which was updated in 2019 with the integration of MP MRI (Table 3.2).

Table 3.2 NICE 2019 prostate cancer investigation and treatment guidelines, AS protocol.

Timing	Tests[a]
At enrolment in active surveillance	Multipara metric MRI if not previously performed
Year 1 of active surveillance	Every 3–4 months: measure PSA[b] Throughout active surveillance: monitor PSA kinetics DRE At 12 months At 12 to 18 months multiparametric MRI
Years 2 and every year thereafter until active surveillance ends	Every 3–6 months: measure PSA[b] Throughout active surveillance: monitor PSA kinetics Every 12 months: DRE

[a] If there is concern about clinical or PSA changes at any time during active surveillance, reassess with multiparametric MRI and/or re‑biopsy.

[b] Could be carried out in primary care if there are agreed shared-care protocols and recall systems.

[c] Could include PSA density and velocity.

[d] Should be performed by a healthcare professional with expertise and confidence in performing DRE. In a large UK trial that informed this protocol, DREs were carried out by a urologist or a nurse specialist.

PSA doubling time or velocity of rise per year, core TRUS re biopsy and MP MRI, are all recognised methods of disease follow up with no single accepted indicator for the requirement to treat (Dall'Era et al., 2012). PSA velocity of > 1ng/l/year on the other hand was found to be highly correlated with conversion to radical treatment but this can be taken as a predictor rather than an indicator of disease progression (Selvadurai et al., 2013).

The long term outcome of AS is being studied in many large trials and no long term results beyond (20 years' time) is yet available. The available studies have suggested

overall good outcome. A systematic review of active surveillance (AS) studies estimated that 1/3 of the patients in this pathway will require an intervention after a median of 2.5 years (Dall'Era et al., 2012). In other large studies as PROTECT, AS was randomly compared to radical prostatectomy and radiotherapy over 10 years' time (for localised prostate cancer) for estimating the risk of disease specific mortality. Similar to other large studies (SSPG), the overall risk of cancer related mortality in 10 years is low at 1%. This is despite statistical significance higher rate of metastasis and disease progression in the active surveillance group which reflects the overall low disease specific mortality (Bill-Axelson et al., 2011).

3.5.3 **Hormonal ablation therapy and chemotherapy**

The concept of prostate gland size variation in response to castration goes back to the 17th century based on an observation of the seasonal variation of testicular and prostate size in animals (Hunter, 1837) . Even though that concept was adopted by further studies suggesting its role in treating prostatic pathologies, it was not until 1941 Huggins et.al proved its palliative effects on metastatic CaP with the reduction of serum acid phosphatases (Huggins et al., 1941, Huggins and Hodges, 1941). Huggins was awarded a Nobel Prize for that discovery. Historically castration was achieved by either oral administration of diethylstilboestrol or surgically by orchiectomy or a combination of both (Blackard, 1975)

In 1977 Schally discovered the structure of Luteinizing hormone releasing hormone (LHRH) secreted by the hypothalamus to stimulate the release of the luteinizing hormone (LH) and follicle stimulating hormone (FSH) from the anterior pituitary gland. When this LHRH analogue was given at regular doses, it was found to down regulate the pituitary receptors and subsequently decrease the circulating androgens

(Vilchezmartinez et al., 1979, Sandow et al., 1978, Schally et al., 2000). Their effects on CaP were tested and the long-term outcome was similar to medical and surgical castration (Tolis et al., 1982).

Androgen receptors blockers can provide further disease control alongside of the castration and some of them can be used as a monotherapy due to their central inhibition of the LH hormone secretion. Hormonal therapy in the UK can be categorised into *Androgen withdrawal* (LHRH or bilateral Orchiectomy) and *Androgen receptor blockage* (combined treatment or mono therapy). Their indications differ according to the disease stage. Furthermore they can be classified according to their use to neoadjuvant, adjuvant and palliative.

LHRH analogues are a well-recognised neoadjuvant treatment with radical radiotherapy which can be the primary treatment to localised and locally advanced CaP (Kumar et al., 2006). The same applies when used as adjuvant treatment (androgen antagonist mono therapy can also be used) with Radiotherapy (Tyrrell et al., 2005)).

The above treatment have curative intention but they can also be used in palliative treatment when disease control is indicated as in metastatic disease. The standard treatment in metastatic disease is LHRH and adding androgen receptors blockers from the start, did not significantly improve the overall survival rate (Dalesio et al., 2000, Seidenfeld et al., 1999). Androgen receptor blockers can still be used as mono therapy where they have more preserved sexual function and fewer hot flushes than LHRH, but are less effective treatment for metastatic disease (Seidenfeld et al., 1999, Seidenfeld et al., 2000).

There has been more recent introduction of hormone ablations therapies which have changed the treatment strategy for both metastatic and non metastatic disease.

Abiraterone acetate is one of those addition and it is essentially a selective irreversible inhibitor of an enzyme (CYP17) which facilitate androgen production in various tissue including CaP cells. Two large RCT reported its benefits for patients with castrate resistant metastatic cancer where it is given along with LHRH or surgical castration and prednisolone (Ryan et al., 2015, de Bono et al., 2011). Abiraterone was also studies Systemic Therapy in Advancing or Metastatic Prostate Cancer: Evaluation of Drug Efficacy (STAMPEDE) for patient with hormone naive Pca who have either metastatic disease or high risk locally advanced disease. The results were published in 2017 and its use with androgen deprivation therapy (ADT) in comparison to ADT alone had a 71% relative improvement in time to treatment failure and 37% difference in overall survival (James et al., 2017).

Docetaxel is another palliative treatment for CaP but not essentially an anti-androgen drug. It is a chemotherapy from the Taxane group which inhibits tumour cell division by binding to the tubulin of the cell microtubules. It was also studied in the STAMPEDE trials and its use in metastatic hormone sensitive CaP increased the median survival from 71 to 81 months (James et al., 2016) and it is only recommended by NICE for this group as backed by evidence from other studies (Sweeney et al., 2015, Vale et al., 2016) (NICE, 2019). For the non-metastatic high risk group, indeed there was significant improvement in the disease control but improvement in the overall survival.

Enzalutamide is another member of the androgen deprivation group which acts on the androgen receptor signalling by inhibiting its nuclear translocation. The available evidence from a large study “AFFIRM” (A Study Evaluating the Efficacy and Safety of the Investigational Drug MDV3100) which evaluated the outcome of Enzalutamide in treatment of patients with CaP whom have received one or more chemotherapy regimen one of which included Docetaxel. The interim analysis of this study was very promising

with 37% risk reduction of mortality in comparison to the placebo (hazard ratio for death, 0.63; 95% CI, 0.53 to 0.75) (Scher et al., 2012). Eligible patients from the placebo group were offered Enzalutamide and the final results for the mortality risk reduction was still the same with a prolonged survival by a median of 4.8 months.

3.5.4 **Brachytherapy**

The use of Brachytherapy for CaP goes back to early 20th century. Minet in 1909 described a technique where a radium tube was placed in a catheter to radiate the prostate. Over the years, the mode of radium delivery has been modified to rectal, perineum and open bladder technique. The general interest in brachytherapy declined with the introduction and development of hormones ablation therapy in the 1970s (Denmeade and Isaacs, 2004). Whitmore (Whitmore et al., 1972) presented an open retro pubic insertion technique, which was criticised for having high complication and recurrence rate due to unequal radiation dose distribution. Holm in 1983 described a trans perineum approach guided by trans rectal ultra sound scan (TRUS), which had a better outcome (Holm et al., 1983). That is the fundamental concept of brachytherapy which can now be categorised into Low Dose (LDR) and high dose rate (HDR). The former, seeds (Iodine-125 (I-125) (most common), Palladium-103 (Pd-103) or Cesium-131 isotopes) are implanted in the prostatic tissue and the radiation dose are delivered over a long period of time. It has less pelvic radiation toxicity to other pelvic organs in comparison to external beam radiotherapy but can have its limitations due to radiation effects on carers. NICE and EAU have both recommended Brachytherapy for localised low risk disease (EAU for a selective group intermediate) as a monotherapy with curative intentions. The evidence from systematic reviews of observational studies demonstrating similar survival rates in low risk localised disease after brachytherapy when compared to other forms of radical treatment (radical prostatectomy & EBRT).

(Doust et al., 2004, Hummel et al., 2003, Nilsson et al., 2004). It has its role in intermediate and high risk disease but only in combination with EBRT. This recommendation were based on the results from the ASCENDE-RT Trial were the combination of low dose brachytherapy and 46 GY EBRT provided better biochemical progression free survival (83%) in comparison to the conventional full dose of 78 GY (62%) (Morris et al., 2017).This does come with a high acute and chronic genitourinary complications in the LDR EBRT combination group (Rodda et al., 2017).

The high dose rate (HDR) is dissimilar to LDR as the implants are temporary and usually introduced through needles or catheters. Their fractions can be delivered in a single or multiple doses and as LDR can be used as monotherapy or in combination with EBRT. EAU is recommending HDR in combination with EBRT for localised disease due to the improved disease free survival with no increase in the toxicity profile when compared to EBRT alone (Hoskin et al., 2012). This lead to an increase in the use of HDR but more evidence is required to support its use as monotherapy for localised disease (Hoskin et al., 2013).

3.5.5 **Radical radiotherapy**

Historically, external radiation therapy was only used to treat bone pain from bone metastases (Denmeade and Isaacs, 2004). The development of the megavoltage linear accelerator at Stanford University in 1950s stimulated research on its use in treating prostate cancer with very promising results (Bagshaw et al., 1965, Bagshaw et al., 1975). Despite its outcome which was comparable to surgical treatment, the older EBRT did not spare healthy pelvic tissue and had significant rate of late complications (Pilepich et al., 1987). The introduction of computerised tomography, allowed mapping the therapy to a three-dimensional image of the target area, leading to better planning, improved

outcome (Ramakrishna and DeWeese, 2001) and reduced gastrointestinal and urinary toxicity over previous techniques (Dearnaley et al., 1999). This 3D conformal radiotherapy remained the preferred choice for EBRT delivery before further enhancement of its technique to deliver intensity modulated radiotherapy (IMRT). The later deliver the radiation in more dynamic way in order to reduce radiation to rectum and bladder and it lead to a significant reduction of late toxicity (Zelefsky et al., 2008).

The conventional dose recommended by NICE and EAU is 74 Gy in 37 fractions which was found to have a better biochemical outcome than lower doses (Heemsbergen et al., 2014). More recent studies reported favourable outcomes with lower doses but with lower fraction (hypofractionation) therefore more Gy per fraction. Conventional versus hypofractionated high-dose intensity-modulated radiotherapy for prostate cancer (CHHiP) trial reported a non inferior outcome in the group treated with 60 Gy in 20 fractions when compared to the conventional dose of 74 Gy in 37 fractions (Dearnaley et al., 2016). The hypofractionated dose is now a standard recommendation in NICE guidelines for localised disease and experimental in the EAU guidelines due to provisional reports suggesting its higher rates of pelvic toxicity (Mottet et al., 2017).

Radiotherapy has a role in treating CaP at all its disease stages. For localised low risk disease, it can be offered as immediate or a deferred treatment option where it can be given as monotherapy (NICE, 2019). This is the only group which did not benefit from adjuvant or neo adjuvant ADT as it did not influence the overall survival (Pilepich et al., 2005). In contrast, neoadjuvant ADT is a standard treatment with IMRT for intermediate and high risk group (NICE, 2019). Their introduction improved both biochemical recurrence and disease free survival with a recommended duration of six months before, during and after completion of treatment in intermediate risk and 3 years after in the high risk group (Bolla et al., 2016, Zapatero et al., 2015). For those 2 groups,

EBRT can be combined with HDR which has reported improvement in disease free survival therefore is a recommended treatment by both EAU and NICE.

For locally advanced disease, EBRT with neoadjuvant ADT for 3 years is a recommended treatment by NICE and EAU with evidence from large RCTs with long term outcomes (Mason et al., 2015, Heemsbergen et al., 2014, Pilepich et al., 2005). It can also be combined with surgery with positive margins on histology which can be given immediately or delayed when PSA reach 0.5 according to EAU guidelines (Fossati et al., 2016). NICE agree with the later as only one study suggested improvement in the overall survival for immediate post-surgery salvage treatment (Thompson et al., 2009).

In palliative treatment, is mainly used for bone metastases to provide symptom control and treat impending spinal cord compression.

3.5.6 **Radical prostatectomy**

Radical prostatectomy is one of the oldest surgical treatment for any cancer. The perineal transvesical approach was described and used in the 17th century to treat urinary obstruction by removing hard tumour from the base of the bladder (Holzbeierlein et al., 2003). This approach was refined to a transperineal approach including removal of Denovillier's fascia, the seminal vesicles, ampullae of vas and part of the trigone, which remained the standard approach for few several decades (Young, 1903). In 1945, Millin presented his retro pubic technique, which involved enucleating the prostate gland leaving the capsule behind. Later he modified it to radical prostatectomy, which unlike the perineal approach allowed lymph node excision and became the gold standard approach for radical prostatectomy (Millin, 1945). This had high rates of blood loss, erectile dysfunction and incontinence and Walsh later refined

this approach with better functional outcome (Walsh et al., 1983). In 1991 Schuessler reported a new intraperitoneal laparoscopic approach for performing the RP which provided and alternative management to the standard open technique but did not have any advantage in the oncological treatment or the complications rates (Schuessler et al., 1997). Not long after in 1999, the era of the Robotic assisted radical prostatectomy (RARP), began with a report on 10 cases performed by J. Binder and W. Kramer (Binder and Kramer, 2001).Similar to Laparoscopic approach RARP did not have oncological outcome advantage and in fact the three approaches had comparable outcome in a large meta analysis (Novara et al., 2012). The same meta analysis estimated better blood transfusion, continence and potency rates which could be attributed to 3 D visualisation and nerve sparing approach of RARP. This was not supported by a Cochrane review which did not report any advantage of RARP apart from lower transfusion rates, less pain and hospital days (Ilic et al., 2017).

A health technology assessment of RARP estimated its cost to be higher than laparoscopic prostatectomy but had a modest increase of the quality-adjusted life-years (Close et al., 2013). This was translated in RARP being more cost effective as long as the unit perform more than 150 cases per year (Close et al., 2013).

This study was cited by Specialised Commissioning Team, NHS England in their published policy for RARP commissioning. They reported an increase in number of RP reported by Hospital Episode Statistics (HES) data 5271 RP in 2013 almost half of those were RARP (Specialised Commissioning Team, 2015). This was predicted to keep rising and most recent national data published by BAUS in 2017 (which was matched with HES data) reported a substantial increase to 7673 cases which were predominantly RARP (6095) (BAUS, 2017). This would align with previous reports of rapidly rising trends of RP with a 10 fold increase between 1991 – 1999 (Oliver et al., 2003).

In the UK it is indicated for localised prostate cancer where there is realistic chance of long term disease control (NICE, 2014). For low risk disease, it can be offered immediate or deferred treatment based on evidence estimating a reduction in mortality of 40% in young patients (<65 years old) with localised low risk disease; number needed to treat being 7 patients to avert one death (Bill-Axelson et al., 2011). This might contradict results from other studies mentioned in the AS section as PROTECT, but a recently published 23 years results by SPCG trial still suggesting a favourable outcome of surgery (Bill-Axelson et al., 2018).

In the intermediate and high risk group, RP can be offered as a radical treatment (NICE) but for the later, it is likely to be combined with other treatment due to the risk of progression (EAU). Surgery was not a favourable radical treatment for locally advanced disease but results from a large study of 1360 patients reported cancer specific survival above 85% at 10 years (Joniau et al., 2015).

3.5.7 Focal treatment

This therapy incudes focal ablation of the localised tumour or the entire gland by cryotherapy or high-intensity focused ultrasound. NICE and EAU consider focal ablation to be experimental and only for localised disease due to the absence of long term outcome. The available data as the systematic review performed by Ramsay et .al, demonstrated low quality data to influence current practice and more well formed RCTs are required to quantify their long term outcome (Ramsay et al., 2015).

3.6 Post treatment follow up

Following disease stabilisation, patients are reviewed with the aims of detecting disease progression at a treatable stage, identifying and treating complications and addressing

any psychological concerns (Rose and Watson 2009). Serum PSA monitoring is a standard among all prostate cancer follow up guidelines following all treatments. However, there is no global accepted single definition for biochemical failure (BCF) for each individual treatment and there is a degree of variability between guidelines (McIntosh et al., 2009). For radical treatment, NICE recommends "*Check PSA levels no earlier than 6 weeks after treatment, at least every 6 months for the first 2 years, and then at least once a year after that".* This is initiated in the cancer care centre then

"After at least 6 months' initial follow‑up, consider a remote follow‑up strategy for people with a stable PSA who have had no significant treatment complications, unless they are taking part in a clinical trial that needs formal clinic-based follow‑up."(NICE, 2019).

Patients on the watchful waiting can be discharged to the primary care *"only if protocols for this have been agreed between the local urological cancer MDT and the relevant primary care organisation(s). Measure their PSA at least once a year" (NICE, 2019).*

3.6.1 **Biochemical Failure definitions**

3.6.1.1.1 Post radical operative prostatectomy

Following RRP, the PSA will be almost undetectable in the serum and two popular and valid definitions for disease spread or recurrence have been developed as alternatives. The first one was reported by Stephenson et.al following the evaluation of 10 definitions of BCF on 3,125 patients post RRP and proposed '*PSA of 0.4 ng/l and risin*g' to be consistent of metastatic disease progression (Stephenson et al., 2006). A further review of the literature by Cookson et al. suggested a definition of '*PSA 0.2 ng/l'* which was discussed and endorsed by American Society for Radiation Oncology (ASTRO)

(Cookson et al., 2007). NICE presented both definitions in their guidelines and suggested a series of PSA levels to discriminate disease recurrence and metastasis from benign tissue remnant. (Grosan and Abraham, 2011)

Further PSA kinetics as *'PSA doubling time less than 3 months'* is associated with disease progression and prostate cancer specific mortality (D'Amico et al., 2004, Freedland et al., 2005).

3.6.1.1.2 Post radiotherapy:

The PSA does not disappear completely following radiation therapy since not all the prostatic tissue is eradicated and it usually takes 2-3 years to reach its nadir (Cookson et al., 2007, Bellera et al., 2009). Patients who had brachytherapy may experience a bounce in the PSA level at a median of 17 months after the implant, which could be falsely interpreted as a BCF (Kuban et al., 2006).

An early definition for BCF following radiation therapy was '*three consecutive PSA rises following the nadir'* (ASTRO, 1997). This definition has been subjected to further analysis and '*nadir + 2 ng/l'* was found to be more sensitive test for BCF following both EBRT (Horwitz et al., 2005, Bellera et al., 2009) and brachytherapy (Kuban et al., 2006).

3.6.1.1.3 Patients on watchful waiting and hormonal ablation therapy:

These pathways are delivered with no curative intention; thus patients will have measurable serum PSA and BCF does not have a strict definition as for post radical treatment. Thus the focus of the follow up in these groups is on the clinical symptoms, DRE, serum testosterone, and assessment of treatment response and side effects (EAU et al., 2015).

3.7 Service organisation of follow up

By default, patients with CaP will have their initial follow up at the cancer care centre until the disease is stabilised. According to NICE, patients with stable disease should be discharged to community follow up with agreed management protocol with their local MDT .There is no information from the literature on the availability of those follow up models or evaluation of their outcome and cost effectiveness apart from 2 RCTs reporting favourable outcome of nurse lead follow up (Faithfull et al., 2001, Helgesen et al., 2000).

More information on the follow up service provision was found in the 2014 National Prostate Cancer Audit (NCAP) annual report which identified five different models in the prostate cancer follow up: Consultant-led clinic, Cancer Nurse Specialist (CNS)-led clinic, Telephone clinic, Community-based specialist follow up and Radiographer-led clinic (only for radiotherapy pathway) (NCPA, 2014). The CNS led was the most prevalent with 77% of the specialist MDTs supporting this model and only 10% have a specialist led community based model.

3.8 Discussion

This chapter presented the declarative (factual) knowledge from the scientific literature and international guidelines on managing CaP. There is an increased incidence of CaP and advancing multimodal diagnostic and treatment tools, which has led to an increase in number of cases diagnosed per year and more cancer survivors. Evidently this has increased the burden on Urology health care with recommendation from NICE for discharging patients with stable disease to the community. This recommendation was not driven by evidence as the available studies are in favour of CNS led follow up (Helgesen et al., 2000, Faithfull et al., 2001). Nevertheless, community follow up should have a more significant role in this domain where the opposite was demonstrated in the

NSCP 2014 (NCPA, 2014). NICE suggested MDT agreed community protocols for enhancing decision making and community follow up. This is evidently lacking and a survey by Watson et. al highlighted the primary care's concern on manging this group of patients in the community due to the lack of the required expertise (Watson et al., 2011). This gap requires more research in CaP decision as conducted in this thesis to develop ES for decision making. Their application can also improve patients 'care as demonstrated by systematic reviews of the computer decision aids outcome (Kawamoto et al., 2005, Garg et al., 2005).

In chapter 2, there was paucity of ES supporting CaP follow up while most systems focused on diagnosis and treatment outcome. This would correlate the evolution of scientific evidence as demonstrated in this chapter where the scope of the large studies is the diagnosis and treatment with no scope on stable disease follow up. This can challenge the development of ES in stable prostate cancer due to the lack of evidence and agreement on some aspects as the frequency of the follow up and timing of PSA. For those grey areas, the developer will rely on expert opinion to generate decision rules. Other parameters that have more solid evidence and census as biochemical failure definition of EBRT or post surgical treatment, can have a straightforward path for conversion into decision rules. The degree of inconsistencies in the later is not high among different international guidelines however the rest have significant variation (McIntosh et al., 2009). McIntosh et al reviewed CaP international guidelines and scored them against a validated appraisal tool "AGREE" and NICE scored the highest reflecting the cost effective approach of its development. Therefore, the declarative knowledge for our ES was primarily driven from NICE guidelines but their compliance where checked against EAU guidelines and the backing evidence were appraised for validity.

3.9 Conclusion

Prostate cancer has a rising impact on the health care and safe cost effective models of care delivery is of high demand. Expert systems are capable decision making tools which can support the development of community based follow up with knowledge from guidelines and domain experts. There is abundant and advancing evidence on disease diagnosis and treatment which is scarce for the stable disease follow up. Subsequently declarative knowledge can be elicited for limited aspects of the stable disease management and further studies are required for procedural knowledge acquisition from the domain experts.

Chapter four - Prostate cancer knowledge

Acquisition (procedural knowledge)

4.1 Synopsis

4.1.1 Background

We aimed to acquire declarative knowledge from Urology experts in order to develop a knowledge base for managing patients with CaP. There was a predicted challenge to elicit this type of knowledge due to the limited ability of experts to access and express this knowledge. Thus, we have developed a methodology with mixed models of knowledge acquisition to acquire the necessary domain knowledge.

4.1.2 Objectives

1. To acquire the declarative knowledge from human experts by running direct unstructured interviews and applying protocol analysis on real cases.
2. Represent the knowledge in decision rules format in order to facilitate the knowledge correction then to structure the knowledge base.
3. Estimate the validity of these fragments by comparing the blinded agreement of the Urology experts with the decision rules using the kappa statistics.

4.1.3 Results

The interviews led to formulation of hand written decision rules. Further testing of these rules on real cases resulted in rules updated in two successive testing cycles (8 then 2 rules) before a 100 % panel consensus was reached. The kappa statistics was estimated and demonstrated overall moderate agreement ($k > 0.7$) on the final version of the decision rules.

4.2 Introduction

Knowledge acquisition (KA) is the bottle neck for developing Expert systems and its failure can arrest the development of the whole system (Benbasat and Dhaliwal, 1989) (Pandey and Mishra, 2009). The various methods of KA have been described above.

We aimed to acquire knowledge from Urology experts on managing patients with stable prostate cancer. These expert knowledge fragments along with the 'static knowledge' acquired from EAU, NICE and other scientific evidence were structures on which to build a knowledge base for the domain. We aimed to enhance the process by using real cases and applying protocol analysis techniques to elicit declarative knowledge from the experts. The statistical representation of the agreement would set a benchmark for the hypothesis testing in the following studies. The acquired procedural knowledge in this chapter were elicited by sub analysis of data from a previous large study for clinical investigation of the ES prototype.

4.3 Previous clinical investigation of the prototype

This study was designed by iQudos the software developers in collaboration with Royal Derby Hospital whom owned 50% of the software shares. This was initiated by MR M J Henley who is a Urology consultant in Royal Derby Hospital proposing an ES for managing stable CaP which was endorsed by the hospital and their research and development department. He developed his own company iQudos and appointed an IT developers and other personnel for developing and managing this clinical investigation study. He acted as the knowledge engineer acquiring the domain knowledge from a local departmental document on CaP management then run unstructured interviews with his

Urology colleague to acquire knowledge for the ES prototype. This prototype was later developed by iQudos into an ES and was believed to be an advanced and would stand rigorous testing in this clinical investigation. This was designed before learning the life cycle of ES as reported in chapter 1 and did not account for mandatory steps as the verification and validation. The end result of this investigation was supposed to support the system registration with the MHRA (Figure 4.1).

The data for this study were all acquired from an existing CNS led follow up clinic and the inclusion and exclusion criteria will be discussed with more details in the methods section. The study compared recommendation on managing those patients with 3 Urologist whom were independent from the software development. The study was supposed to run until 100% agreement was reached among all parties with 100 cases in each of the first two rounds then 51 in the last one. The study had the North London Ethical committee approval on the 18th of October 2010 (REC 10/H0724/63) and the protocol is in appendix 1.

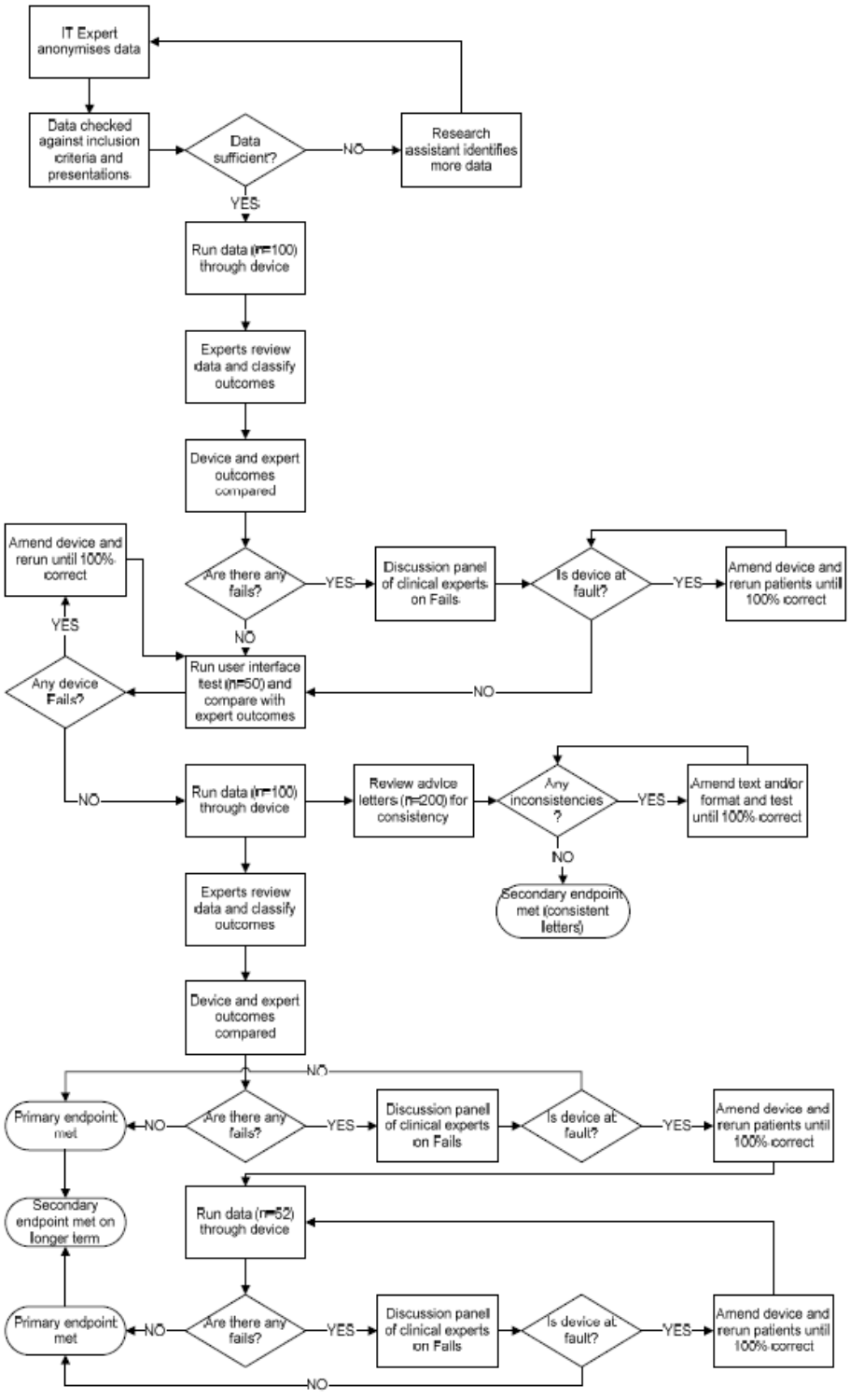


Figure 4.1 Flow chart for the clinical investigation plan

The study run 3 rounds of case analysis followed by a panel meeting to discuss discrepancies. The main aim was to correct thee system until 100% accurate and compliant with the investigators' recommendations.

4.4 Outcome of the clinical investigation

The study completed the full cycle with 251 cases being included and analysed by the software and the clinical investigators in 3 rounds as per the study flow chart. The results of each round are discussed below in the results section and the agreement was fair in the first round and it improved until 100 % agreement was reached in the last round of 51 cases. This was believed to be the primary end point of the study and further analysis was planned to meet the secondary end point which was predominantly consistency checking.

The study monitoring was performed at that stage and reported significant concerns. The main concerns were deviation from the study protocol and poor data keeping which led to withdrawal of the research and development department support and termination of the clinical investigation despite completing the full cycle. The author of this thesis was one of the clinical investigators and have summarised below the key elements that have led to the poor conduct of that study despite the clinicians and developers efforts to refine the system.

The first challenge was the complexity of the protocol where the clinical investigators had to codify the clinical information and outcome as seen in the mind map in (Figure 4.2). For instance, patient presenting with LUTS and ED with high creatine, high Alk Phos and stable PSA post radiotherapy were codified as B1 HS1 HS2 ST2 ST3 then recorded in the outcome sheet. This has introduced an unnecessary layer of false errors due to applying the wrong code. This was clear in the first round of 100 cases therefore, in the second round, the investigators had a chance to review the software codes before the panel meeting to choose the cases with true disagreement. This was thought to be

deviation from the protocol but on reading the small prints of the protocol, this was still allowed but it was felt that the results are still not representative of the true agreement.

Second limiting matter was the absence of the clinical monitor which had led to poor data keeping. This has hindered all attempts for salvaging the study by refining the protocol and resubmitting to REC.

Furthermore, the study was designed as pass or fail per case not per decision rule while the latter is what will have the error and requirement for correction. This will be discussed and rectified in the testing studies in chapter 5 and 6.

The study did not apply weights on the disagreements, therefore a disagreement on managing erectly dysfunction will have the same weight as missing cancer progression. Clearly both should be discriminated and accounted for in the validation.

Despite the above, the study has run the full course and the clinical investigators have indeed solved the cases then discussed their management in a panel meeting chaired by an independent Urologist. Those discussion led to the software update and the last 51 cases were reviewed and recorded according to the protocol and the agreement was 100%.

It was concluded that the data from this study should not be used for verification or validation which were performed later in chapter 5 and 6. However, this study has led to acquiring more knowledge and upgrade of the prototype to a presumed valid ES. Therefore, data sub analysis was performed on this study to report the acquired knowledge and be a platform for further validation studies.

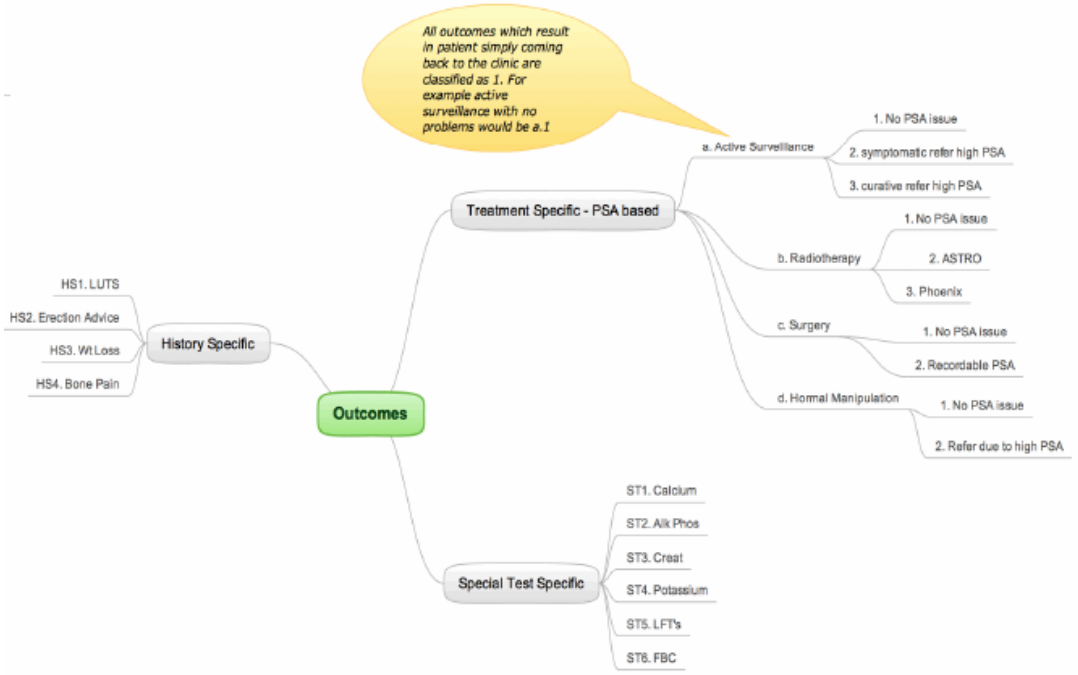


Figure 4.2 Mind map for coding the cases input and outcome
For example: patient presenting with LUTS and ED and have high creatine and Alk Phos with stable PSA post radiotherapy: B1 HS1 HS2 ST2 ST3.

4.5 **Methods**

The static knowledge was derived from NICE and EAU guidelines as presented in chapter 3. The main focus in this study was on the acquiring the procedural aspect of the knowledge from human experts. As mentioned before, Mr Henley who acted as the knowledge engineer run unstructured interviews to discuss and refine the domain knowledge. He interviewed three consultants with different training background and Urological interest. The definition of expert in this domain is anyone with qualification to review and advice on the management of the majority of this group of patients independently. This would include Urology Consultants, senior Urology specialist registrars and CNS specialists. The unstructured interviews were manually recorded by the knowledge engineer focusing on an existing local pathway for the follow up of patients with CaP. This was documented as remarks on the department pathway and

lead to the decision rules and mind maps which are presented in appendix 1 & 2. This was then applied in a rule based format and used to build the prototype for the clinical investigation.

Watchful Waiting rules

Refer if PSA > 20.

If [∂ PSA + current PSA > PSA max]. In other words, previous rise in PSA added to current PSA greater than threshold (levels to be set by Urologist or 20 by default) then see at 3 months rather than 6. (∂ = the rise in PSA i.e. current – previous PSA)

Figure 4.3 Decision Rule example
The above handwritten rule is for detecting biochemical failure for patients on the watchful waiting pathway. They are either referred if PSA > 20 or seen more frequent if there is a rising trend. Alternatively, the patients can be placed on consultant specific pathway with different PSA thresholds.

The handwritten decision rules (DR) were used to acquire more knowledge from the clinical investigators by applying them on real cases. The cases were included from an existing database of patients' follow up in a Teaching hospital run by Urology cancer nurse specialists. This cohort of patients have been diagnosed and treated for CaP and been previously discussed in the Urology MDT. Those were later discharged from the consultant led clinic in Royal Derby Hospital to this CNS clinic. There is not specific criteria for patients' eligibility to be managed in this clinic as the CNS are well trained and directly linked to the local MDT and the referring consultant. They tend to manage patients with more stable disease than newly diagnosed cases and or having active progression.

Data from this clinic were anonymised by the database manager who is also the software developer, Mr Steven Pick then presented to study clinical monitor for examination against the inclusion criteria (Figure 4.4). The clinic letter generated for each visit

usually have sufficient information on the case allowing a decision to be made against the inclusion criteria.

4.5.1 Cases inclusion

4.5.1.1 Eligibility criteria

All adult patients seen in this clinic with known prostate cancer whom had their primary treatment assigned and delivered. All risk groups and treatment pathways are eligible for review against inclusion criteria.

4.5.1.2 Inclusion criteria

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

- Known prostate cancer treatment pathway.
- Details on his PSA on presentation and most recent results.
- Protocol also stated that PSA done within the last 60 days prior to the appointment but this was not strictly followed.

With or without (optional requirement):

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

4.5.1.3 Exclusion criteria

- Cases with newly diagnosed prostate cancer that are waiting for treatment decision.

- Cases with known prostate cancer that are being evaluated for known recent disease progression.
- Cases with no clinical details of their initial cancer presentation such as Gleason score, PSA, and tumour grade which would influence the follow up decision making.

4.5.2 **Cases preparation**

All included cases were anonymised and their clinical outcome was eliminated. This was performed by the software developer and database manager and included removing all identifiable data and replacing them by case study number. Thus, each case had information on diagnosis (CaP, Gleason score if not clinically diagnosed cancer, PSA) treatment received and current clinical symptoms and blood results (Figure 4.4)

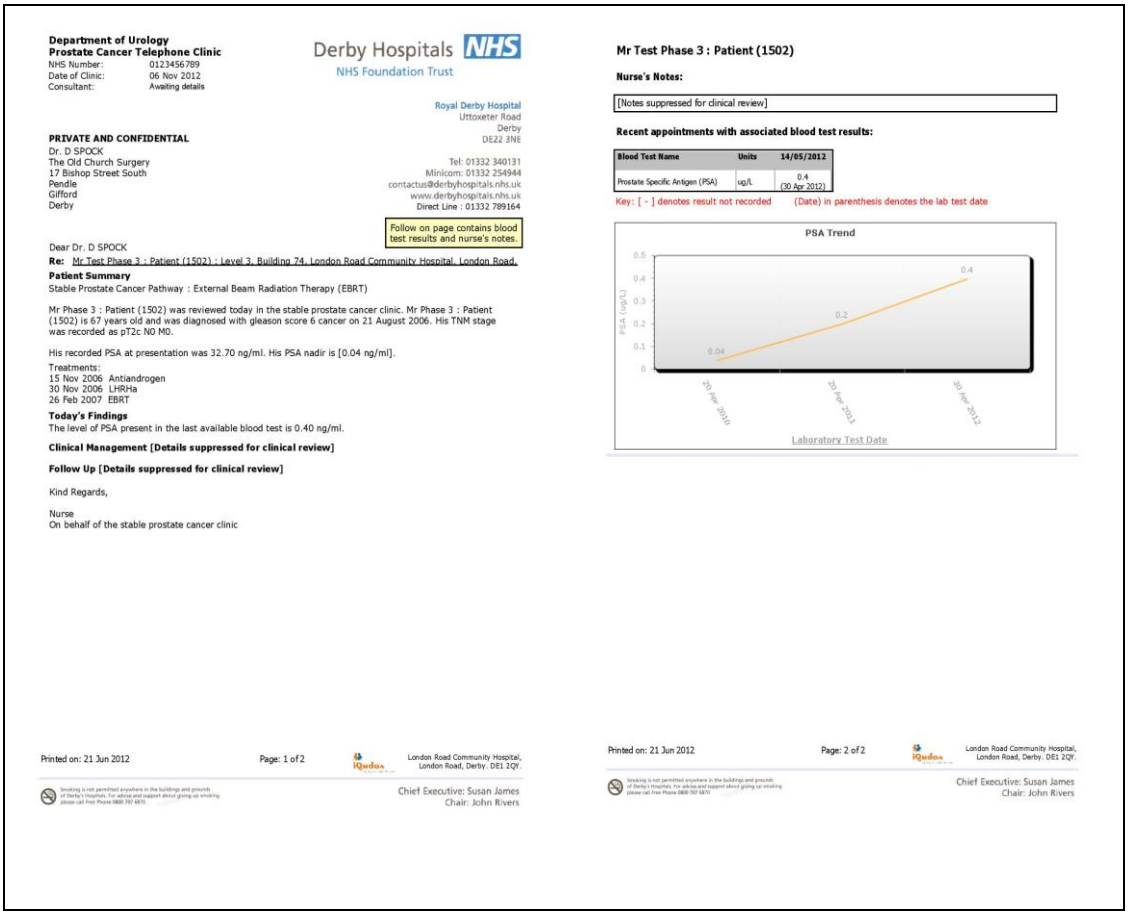
4.5.3 **Case rating**

In the clinical investigation study, included cases (n=251 in total, in three main phases, cohort 1 n=100, cohort 2 n=100 and cohort 3 n=51) were independently reviewed by the clinical investigators (INV). The INV were all recruited from the Urology department of the Royal Derby Hospital and were three in each phase. Only one INV was present in the three phases and they were all qualified consultants or registrars (more in the results section). The CNS were not included as they run the follow up may be familiar with the cases. INV were enrolled in this study following the definition of expert in the methods section.

The cases were reviewed and rated independently by the INV according to the written decision rules (DR) then codified following the mind map (Figure 4.3). Similarly, the prototype assigned ratings to the same case by rules driven from the DR then all rating from the investigators and the ES were compared for discrepancies. The cases with

discrepancies where then included for discussion in the panel meeting. The panel included the knowledge engineer, database manager, clinical monitor, 3 INV and a urology consultant as a chair.

According to the original protocol of the clinical investigation, a cohort of 100 patients was reviewed and the threshold of 100 % agreement was not reached. Therefore, another cohort of 100 patients were reviewed following the same methodology which had better agreement but did not reach the 100% cut off. A last set of 51 cases achieved that threshold and the primary endpoint was then reached. The study was powered at 80%, a 95% Confidence Interval of 0.3% error rate with precision +/-1.5%. This presumed error in the first 2 rounds and the use of a final set of 52 cases.



Department of Urology
Prostate Cancer Telephone Clinic
NHS Number: 0123456789
Date of Clinic: 06 Nov 2012
Consultant: Awaiting details

Derby Hospitals NHS
NHS Foundation Trust

Royal Derby Hospital
Uttoxeter Road
Derby
DE22 3NE

Tel: 01332 340131
Minicom: 01332 254944
contactus@derbyhospitals.nhs.uk
www.derbyhospitals.nhs.uk
Direct Line : 01332 789164

PRIVATE AND CONFIDENTIAL
Dr. D SPOCK
The Old Church Surgery
17 Bishop Street South
Pendle
Gifford
Derby

Follow on page contains blood test results and nurse's notes.

Dear Dr. D SPOCK

Re: Mr Test Phase 3 : Patient (1502) : Level 3, Building 74, London Road Community Hospital, London Road.

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

Mr Phase 3 : Patient (1502) was reviewed today in the stable prostate cancer clinic. Mr Phase 3 : Patient (1502) is 67 years old and was diagnosed with gleason score 6 cancer on 21 August 2006. His TNM stage was recorded as pT2c N0 M0.

His recorded PSA at presentation was 32.70 ng/ml. His PSA nadir is [0.04 ng/ml].

Treatments:
15 Nov 2006 Antiandrogen
30 Nov 2006 LHRHa
26 Feb 2007 EBRT

Today's Findings
The level of PSA present in the last available blood test is 0.40 ng/ml.

Clinical Management [Details suppressed for clinical review]

Follow Up [Details suppressed for clinical review]

Kind Regards,

Nurse
On behalf of the stable prostate cancer clinic

Printed on: 21 Jun 2012 Page: 1 of 2 London Road Community Hospital, London Road, Derby. DE1 2QY.

Chief Executive: Susan James
Chair: John Rivers

Mr Test Phase 3 : Patient (1502)

Nurse's Notes:

[Notes suppressed for clinical review]

Recent appointments with associated blood test results:

Blood Test Name	Units	14/05/2012
Prostate Specific Antigen (PSA)	ug/L	0.4 (30 Apr 2012)

Key: [-] denotes result not recorded (Date) in parenthesis denotes the lab test date

Printed on: 21 Jun 2012 Page: 2 of 2 London Road Community Hospital, London Road, Derby. DE1 2QY.

Chief Executive: Susan James
Chair: John Rivers

Figure 4.4 A sample of the cases before rating by the investigators
Real cases were included from an existing database of real patients follow up service. All cases were anonymised and had their real clinical management supressed before being presented to the investigators for rating

4.5.4 **Sub analysis methodology**

The main aim of this sub analysis is analysing the KA and representing this from the data of the clinical investigation. For this purpose, the study methodology was simplified to compare the data solved by the INV and comparing this to the cases solved by the decision rules. Therefore, the prototype was not part of the comparison and substituted with the handwritten decision rules as the software performance is not relevant to the knowledge gained from the experts. Furthermore, the data required for this sub analysis is objective and well documented and avoids the confusion introduced from the mind map coding.

A methodology was required to represent the panel meetings discussion and "protocol analysis" was felt to be the most representative model of this process. Protocol analysis can be defined as 'processing and transformation of texts gained by transcribing protocols of loud thinking during a problem-solving process' (Diederich et al., 1987). The think aloud and cross examination protocol methods were applied in the panel meeting to discuss discrepancies and update the DR. This resulted in adding, substituting or amending the DR and as per the initial protocol, any update would require a retest with a further cohort of 100 cases then 51 cases until 100% agreement was reached. Therefore, the flow of the original study was appropriate for this sub analysis and no modifications were required.

4.5.5 **Statistical analysis**

The initial study suggested the kappa (k) statistics to estimate the agreement between the INV and the prototype (Cohen, 1960). The level of agreement was not decided following Altman (1990) but aimed for 100% for passing the device.

Even though the sub analysis primary objective was the knowledge acquisition and the protocol analysis. The kappa statistics was still estimated for the 3 rounds as per the initial protocol but with deviation from validating the software to refining the DR. Therefore, reaching the 100% agreement was the threshold for stopping the KA rather than the point of validity. Furthermore, the statistical results were used to guide powering further validations studies and hypothesis testing (Figure 4.5).

The protocol analysis was performed on cases where the investigators disagreed with the DR and proposed a different management plan which was later discussed in the panel meeting. The panel discussion evaluated the rules suggestions resulted from the protocol analysis (applying the think aloud and cross examination techniques) and analysed them against the existing decision rules. The result of this analysis was either pass or fail for each rule (of the decision rules). The failing rules were either removed, amended or substituted. This process was statistical represented by objective numbers and percentages.

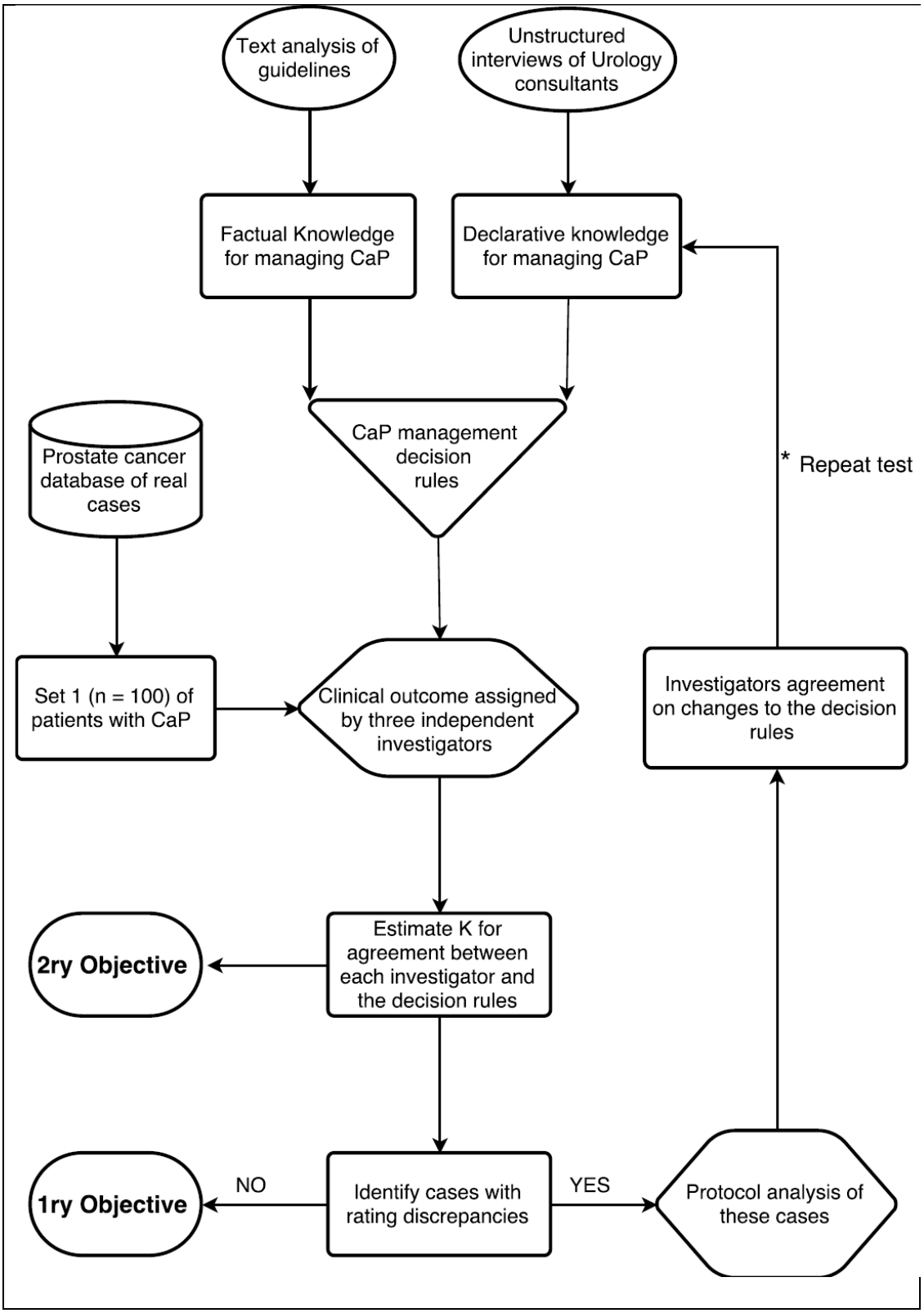


Figure 4.5 Study flow chart

Flow chart demonstrating the whole study progress to acquire knowledge for managing patients with stable prostate cancer (CaP). Primary objective will be reached when 100% agreement is reached and knowledge acquisition is complete.

4.6 Results

The first cohort of 100 hundred cases were initially included and rated by the 3 investigators independently (two consultants and one senior specialist registrar). The blinded agreement between INVs and the DR were calculated by kappa statistics to demonstrate moderate agreement (INV1 k=0.59, INV2 k=0.53, INV3 k=0.63 $p < 0.005$) (Table 4.2). The cases with discrepancies ($n > 50$) between the INV and DR were analysed in the panel meeting. The majority of the discrepancies were due to coding error rather than true disagreement. The INV performed protocol analysis on the cases with true discrepancies and the knowledge engineer acquired procedural knowledge from this process then represented it in the form of decision rules upgrade (n=8). This led to upgrade of 8 rules with 4 rules substitution, 3 rules amendment, 1 rule addition and no rules were deleted (Table 4.1).

Rule	Before	After
Bone pain and high alkaline phosphatase	Advise to GP	Refer to hospital
Missing PSA	Advise GP	Repeat and review
Missing Blood	Repeat and review	Advise GP
Active surveillance	Suitable for community	Not suitable for community
Isolated high ALP	Consider mets	Change wording to other differential diagnosis
Erectly dysfunction	Advise GP	Include use of medications

Active surveillance (with no intention to cure)	Active surveillance	Watchful waiting
Rapid rise in PSA		Repeat PSA and urgent follow up

Table 4.1 Decision rules change following the first panel meeting
The panel discussed more than 50 real cases managed with the decision rules and their discussion led to 8 rules update. The top 4 had their outcome changed. The three after only had their wording amended and the last one was a new rule.

A second cohort of 100 cases was included and rated by the three INV (same 2 consultants and a different specialist Urology registrar). The kappa coefficient improved in this round and progressed to moderate agreement (INV1 k=0.59, INV2 k=0.62, INV3 k=0.62 $p < 0.005$) (Table 4.3). Similarly, the panel analysed the discrepancies and, on most occasions, the disagreement were due to coding errors and they agreed on most occasions with the ratings of the DR. The protocol analysis resulted in acquiring 2 new rules which were added to the decision rules. The rest of the rules passed without any changes (appendix 2).

A last set of 51 cases was included and analysed by the 3 INV (one consultant was present in the previous two rounds and 2 specialist registrars one new to the study). The agreement was significantly higher in this group (INV1 k=0.74, INV2 k=0.79, INV k=0.71 $p < 0.005$) (Table 4.4). The panel meeting discussed the discrepancies and they agreed with the DR on all occasions. Protocol analysis was not performed for this set and all rules passed without any changes.

INV	KAPPA	SE	P
INV 1	0.59	0.099	< 0.005
INV 2	0.53	0.128	< 0.005
INV 3	0.63	0.107	< 0.005

Table 4.2 Kappa statistics for set 1

INV	KAPPA	SE	P
INV 1	0.59	0.107	< 0.005
INV 2	0.62	0.095	< 0.005
INV 3	0.62	0.113	< 0.005

Table 4.3 Kappa statistics for set 2

INV	KAPPA	SE	P
INV 1	0.74	0.11	< 0.005
INV 2	0.79	0.116	< 0.005
INV 3	0.71	0.115	< 0.005

Table 4.4 Kappa statistics for set 3

4.7 Discussion

This study acquired procedural knowledge for managing CaP from Urology experts. This was combined with the declarative knowledge from chapter 3 for developing the ES prototype. The clinical investigation aimed to test it by comparing human experts' reasoning to the prototype for validation and registration with the MHRA as a standalone software. The study had a complex protocol which the investigators struggled to follow and with lack of monitoring, led to an unfavourable outcome and termination of the study despite completing the three phases.

Nevertheless, this study had three panel meeting with discussion of disagreements and the protocol analysis of the cases, led to acquiring more procedural knowledge and upgrading the prototype. This was sub analysed in this chapter with a salvage study with a scope on the protocol analysis and the knowledge acquired for the software development.

The clinical investigation did not specify the knowledge acquisition as a clear objective as the prototype was thought to be mature and the study was to support its registration

with the MHRA. In retrospect, this would have been challenging even with favourable outcome of the original study due to many reasons. First of all, the testing had a poor definition of the validation and verification in their methods which as discussed before, is a standard phase in the development (OKEEFE). This could have led to shallow testing and poor definition of error correction. Those definitions should have been clarified in the protocol and their requirement for correction. For instance, error in wordings or typos should have a simpler pathway for corrections than inadequate knowledge missing cancer progression. Furthermore, the study tested the whole system and where it should have focused on the decision rules which evidently needed upgrade and correction. This invites the risk of suboptimal testing of the rules as described in chapter 5 and 6.

Combining text analysis of guidelines and scientific evidence appraisal with protocol analysis, is not an uncommon methodology for knowledge acquisition (Diederich et al., 1987). This is not the only model and a Delphi study could have performed the same test and had a simpler and less time consuming study. However, it would not acquire the knowledge gained from the panel meetings and estimate the agreement which set the benchmark for other studies testing the software.

Further limitation of this study is in the cases' diversity and not being a true representative of the domain. This limitation was minimised by recruiting real cases from a secondary care clinic and performing the test in stages until validity and completeness were achieved.

4.8 Conclusion

Procedural knowledge acquisition is a complex and crucial step in the ES development. Applying real cases on the handwritten decision rules allowed gaining more knowledge

from the experts but having a complex protocol led to adverse effect. The handwritten decision rules have simplified the process of error correction and can be developed into an ES knowledge base.

Chapter five - Knowledge validation

5.1 Synopsis

5.1.1 Background

Knowledge validation is a critical stage in developing expert systems. There is lack of evidence and formality of their methodology. We aimed to validate knowledge acquired from Urology experts and scientific evidence for managing patients with CaP. The secondary objective was to represent the knowledge in a valid format facilitating the development of decision support tools.

5.1.2 Methods

The acquired knowledge was reduced into fragments which were then constructed into decision trees. Further data was acquired from an existing database of patients with CaP. The decision trees were applied to these patients' data (n=100) to generate solved cases (cases with management plans from decision trees) then presented to two independent reviewers for rating.

The reviewers' input was quantitatively analysed for validity (kappa statistics) and reliability and qualitatively analysed validity analysis of each rule then for redundancy.

5.1.3 Results

Quantitative analysis estimated substantial agreement between the reviewers on the validity of the decision trees (weighted kappa = 0.96, 95% CI 0.94=0.99) with 100 % reliability. Qualitative analysis found all rules (n = 36) to be acceptable with only two requiring some minor adjustments.

5.2 Introduction

In this study, we aimed to validate the handwritten decision rules by estimating the blinded agreement between Urology experts and the decision rules on retrospective real data gathered from a teaching hospital database. The valid knowledge was represented in decision trees format to facilitate the development of AI knowledge base. The agreement estimated in this study was considered in deciding the statistical values in the hypothesis testing in the live trial of the system.

5.3 Methods

5.3.1 Knowledge design

Expert knowledge generated in chapter 4 resulted in a set of handwritten decision rules which were applied on a mind map in order to develop decision trees (DT) (protocol in appendix 4). These DT were further validated in this study and the result was formatted into a rule driven knowledge base (Figure 5.1).

5.3.2 Case selection, synthesis and analysis

All cases were acquired from an existing database of CaP follow up service in Derby Teaching Hospitals Foundation Trust. In the first instance, a set of audited cases was anonymised and examined against the selection criteria. The study was classed as service evaluation by the University of Nottingham Ethics and approved by the royal Derby hospital research and development department (approval in appendix 3).

5.3.2.1 Eligibility criteria

All adult patients seen in the CNS clinic mentioned in the methods section in chapter 4 with known prostate cancer whom had their primary treatment assigned and delivered.

5.3.2.2 Inclusion criteria

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

- Known prostate cancer treatment pathway.
- Details of PSA on presentation and most recent results.

With or without (optional requirement):

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

5.3.2.3 Exclusion criteria

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

Two hundred and three cases between June 2013 and May 2014 (11 months), were included. The initial inspection excluded two due to missing data. The rest (n = 201) were examined to stratify cases with similar clinical presentations and progression into groups.

The case stratification considered the hierarchy of the clinical variables according to their significance (Table 5.1). The primary defining criterion was the treatment pathway or the most significant treatment in the case of multiple modality treatments. The treatment modality was linked to the current PSA progression and then represented by a code to facilitate grouping of similar cases. For example, all cases on hormonal ablation therapy pathway with PSA indicative of recurrence were given a similar code. These cases were further categorised according to the associated symptoms (lower urinary tract, bone pain, weight loss and erectile dysfunction) and abnormal blood test if present (livers or kidney function, calcium and electrolytes, haemoglobin, potassium and alkaline phosphatase). Each symptom or blood test was also represented by a code. Thus, for each case the codes demonstrated treatment pathway, PSA progression and symptoms. All cases were substituted with these codes then tabulated on a spread sheet (Excel, Microsoft Corp, Seattle) for comparison (Table 5.1). On the excel sheet, all cases with similar codes were placed in the same group and to avoid repetition, only one from each group was included in the study. This was done to avoid replication of identical pathway of the knowledge and test all of its fragments.

For the rare clinical presentations which were not covered by this set of cases, a further set of altered real cases was synthesised and included. This allowed testing the rules covering rare presentation and completing the testing set. Synthetisation was performed by the clinical monitor using real life scenarios where possible to trigger those rare presentations.

Case	Treatment Pathway Hormone ablation = H Radical Prostatectomy Radical radiotherapy Active surveillance Watchful waiting = w Consultant specified = c	PSA progression No BCF = 1 Rise but no BCF = 3 BCF = 4 ASTRO = 4 PHOENIX = 4	Disease complications LUTS = LU Bone pain = BP ED = ED Weight loss = WL	Bone metastasis Bone pain = BP + Calcium disturbance = Ca High ALP = ALP*	Other blood test Creatinine = CR Anaemia = HB Potassium = K Liver Function test = LFT	Final codes *
Case 1	Radical radiotherapy = R	ASTRO = 4	LUTS = LU	None	CR	R4 LU CR
Case 2	Retropubic prostatectomy = RP	NO BCF = 1	Erectile dys = ED	Bone pain = BP	Anaemia = HB	RP1 ED BP HB

Table 5.1 Case stratification

Each case attained one or more code according to the disease history, PSA progression, bone metastasis, complications and other blood test. A final code is produced for each case and cases with similar codes were stratified in the same group. There for all patients in the same category have same codes and representing similar clinical information and disease progression. ** ALP = Alkaline phosphatase serum level*

For internal validation, a set of consecutive real unaltered cases was examined and included to predict the real time accuracy of the decision rules and simulate a pilot trial of the tool.

For all included cases (testing and validation set total = 100), the nurse's recommendations and real time clinic outcome were eliminated then replaced by suggestions derived from the decision trees. These suggestions were all objective outcomes for each case following mapping their clinical presentation with the decision trees recommendation. The case clinical presentation and the decision trees outcome were combined together in a clinic letter (solved case) and the case was then presented to the two independent Urology consultants for case rating. The case evaluation by the investigators (INV) included a quantitative rating of the rules where they assign agree or disagree to the case outcome then qualitative evaluation by leaving feedback in free text box (appendix 5).

The clinical investigators (n=2) were recruited from the same Urology department and they were both independent from the knowledge development team. Therefore they were not familiar with the decision rules from the previous study and their ratings was performed independently.

5.3.3 **Study analysis**

This study applied a mixed methodology design guided by models (across stage mixed model design) as described by Johnson and Onwuegbuzie (2004). The qualitative analysis had the highest weight in this study where each rule in the DT was assigned PASS or FAIL according to investigators disagreement (Figure 5.1). This resulted from the analysis of the investigators' comment on their disagreement with the rules to categorise the outcome for each rule into PASS (all agreement and no negative comments), PASS with possible updates (if only minor disagreements reported) and FAIL (major disagreement and/or negative comments). If a major disagreement was encountered, then a team consensus meeting would decide the rule correction or substitution.

Further qualitative analysis of rules assessed the individual rule adequacy and redundancy by examining the clinical investigators comments.

The quantitative analysis for the validity estimated the true agreement of the independent INV with the rules. This agreement was estimated by the Kappa statistics (Cohen, 1968, Cohen, 1960), IBM SPSS 22 and VassarStats) to assess the true agreement while eliminating the agreement by chance (Altman, 1990). Weighted Kappa values were calculated in order to reflect the level of the disagreement according to the clinical significance.

Tool objective reliability was assessed by comparing the outcomes of cases with similar case presentations. It would be expected that these cases would have the same outcome and the acceptable reliability for this study was 100%.

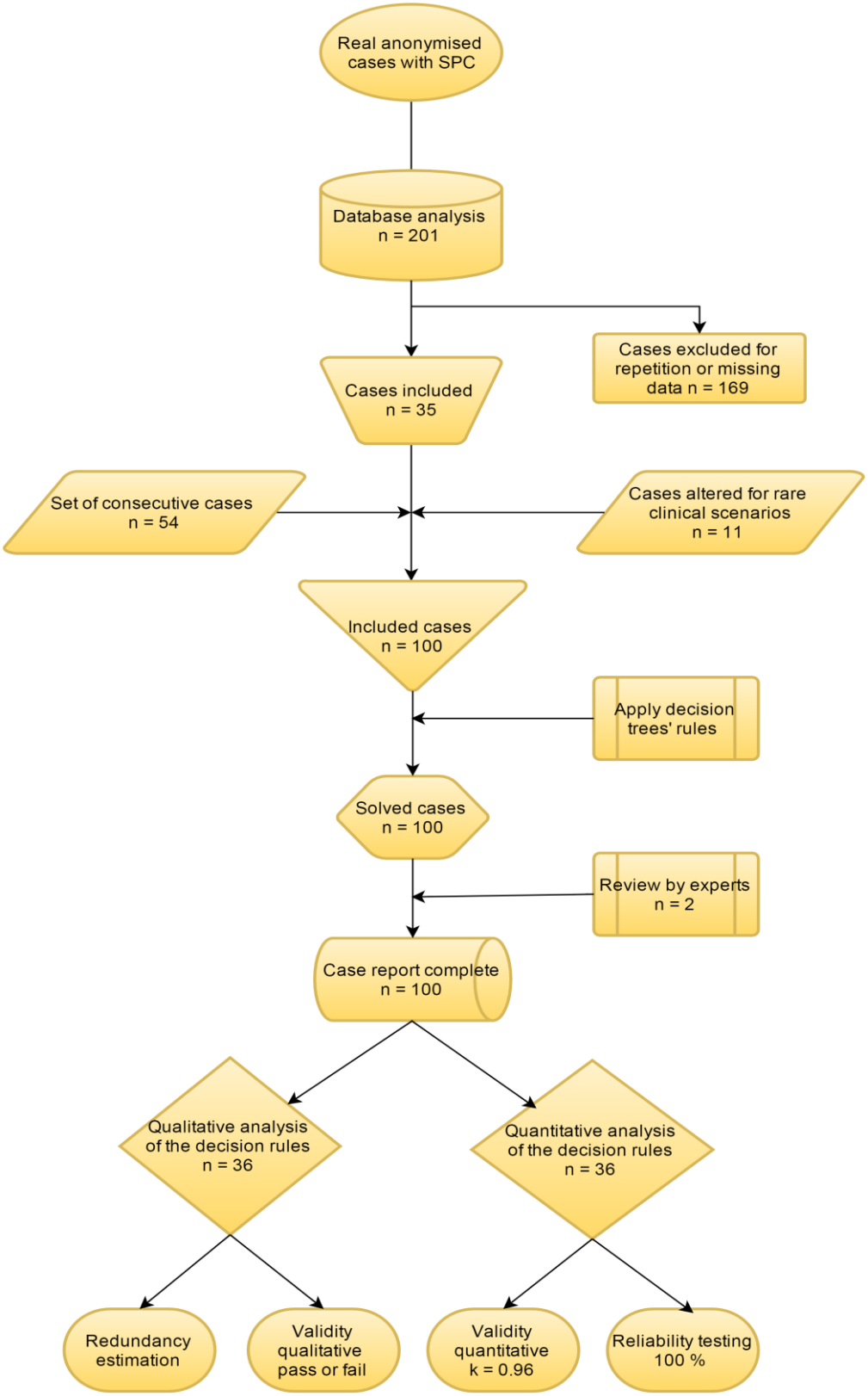


Figure 5.1 Decision trees development and validation

The eligible 201 cases with Stable prostate cancer (SPC), were analysed and only 35 were included. The rest were excluded due to close similarity of the clinical presentation and outcome on the included 35 cases (as stable disease post radical surgery and plan for review in 6 months). Eleven more cases were required to test rare clinical presentation covered by the decision tree to have a total of 46 cases to complete the testing set. A further set of 54 consecutive cases were added for internal validation to have a total of 100 cases. The rules (n = 36) were then quantitatively and qualitatively assessed for validity

5.4 Results

5.4.1 Testing set

The 201 cases were stratified into 35 clinical presentation groups where each group represent cases with same presentation and outcome (Figure 5.2). Therefore only one case was selected from each group as a representative of this clinical scenario (n = 35) to avoid unnecessary repetition as mentioned in the methods (169 excluded). This set of 35 cases tested all the decision rules but 12 (total number of rules = 36) representing rare clinical presentations. This required altering a further set of 11 cases to have a total of 46 cases in the testing set. Thus, each rule was tested at least once by this set.

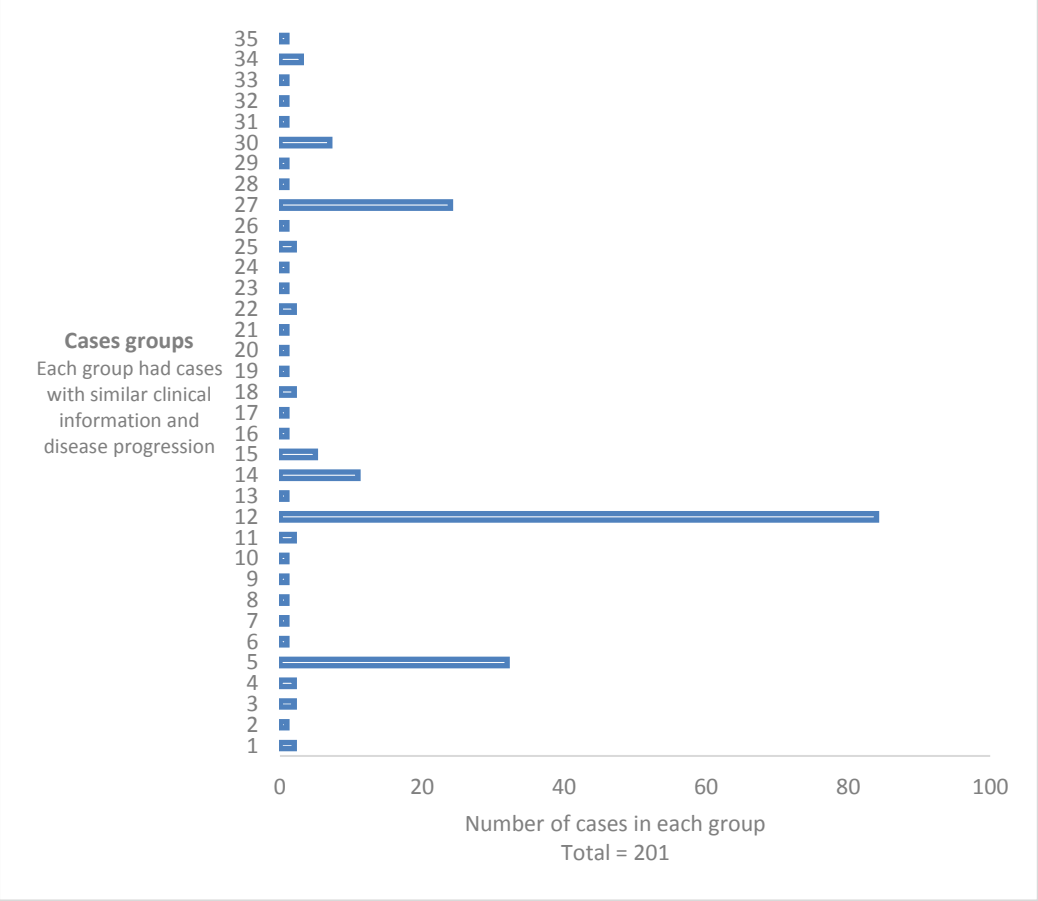


Figure 5.2 Cases classified in groups according to their clinical presentation
Cases (n = 201) with similar codes representing their clinical presentation and disease progression were stratified in the same group to have a total of 35 groups represented on the vertical axis. One case from each group was included in the study to avoid repetition and they tested all rules but twelve where eleven further cases were synthesised to complete the initial test set (n = 46).

5.4.2 **Validation set**

A further set of 55 consecutive real cases was acquired from the same CNS-led clinic database and reviewed. Only one case was excluded (for missing data) and the rest of the 54 cases were included without any alteration for the internal validity. Thus, total of 100 cases (46 in the test set and 54 in the validation set) were included and rated by the DT and the clinical investigators.

The median age of cases was 76 (range 54 - 92) and all had a primary diagnosis of prostate cancer. All cases were stratified in one of six treatment pathways (Figure 5.3): hormone ablation therapy 19 %, radiation therapy 52%, radical retropubic prostatectomy 16%, watchful waiting 7%, active surveillance 2% and consultant specified threshold 4%. The cases classification according to severity at presentation (D'Amico et al., 1998) were 56% high, 31% intermediate and 13% low risk.

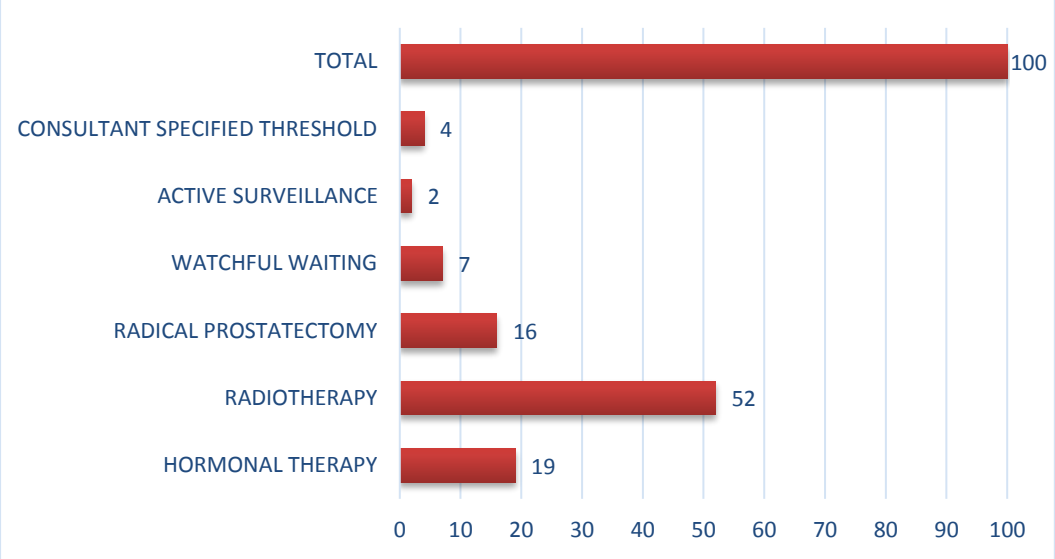


Figure 5.3 Included cases stratification according to the treatment pathway
The testing (n = 46) and validation sets (n = 54) analysis demonstrated prevalence of Radiotherapy in the follow up clinic. Active surveillance had limited presence as they are likely to be closely monitored by the specialist.

5.4.3 **Qualitative analysis**

The decision rules were triggered 168 times for solving all cases clinical presentations. Each case can trigger more than one rule, therefore similar to real life some cases with multiple problems will have multiple management recommendations.

The validation analysis of the 36 rules according to the INV disagreement and team consensus meeting, classified 34 rules as accurate (pass), two rules accurate with minor errors (pass with possible changes) and no rules have failed. Further on, the adequacy analysis was satisfactory for all of the rules and no suggestion for any rules addition. However, the redundancy analysis identified two rules to be redundant if combined for solving the same case (rule 26 and 12).

5.4.4 **Quantitative analysis**

The blinded agreement between both investigators on the decision trees recommendations was substantial with unweighted kappa, k = 0.93 (95 % CI 0.88-0.98), linear weighting k = 0.95 (95% CI 0.91-0.99) and quadratic Weighting k = 0.96 (95% CI 0.94-0.99) (Table 5.2).

The reliability testing demonstrated 100% repeatability. Without exception, cases presenting with similar symptoms and received similar treatment, had the same outcome assigned by the decision tree analysis.

TYPE OF KAPPA	OBSERVED KAPPA	STANDARD ERROR	95% CI
Unweighted	0.93	0.026	0.88-0.98
Linear weighting	0.95	0.019	0.91-0.98
Quadratic weighting	0.96	0.014	0.94-0.99

Table 5.2 Quantitative analysis and kappa statistics estimation
Quadratic kappa was estimated for the blinded agreement of the two investigators over the validity of the decision rules

5.5 Discussion

This study validated a knowledge base for managing patients with CaP that was acquired from written text and human domain experts. This is the first validation model to apply quantitative and qualitative statistical analysis for appraising medical knowledge while providing an objective methodology for identifying and grading errors. Other available appraisal models for knowledge driven decision support tools struggle to integrate heterogeneous knowledge formats from humans and written text and to link the validation process to the clinical outcome (Vanhaecht et al., 2006).

Other knowledge validation methodologies applied by AI systems, aim to reiterate the knowledge content and until their outcome is matching the expert (Okeefe and Oleary, 1993). This approach is limited by the accuracy of the expert decisions and criticised for considering human decisions as gold standard. In this study, we aimed to eliminate this error by applying a mixed methodology in order to quantify the agreement then evaluate individual knowledge fragment qualitatively. Furthermore, the rating was done blindly and the investigators were independent from the knowledge development team.

This study estimated the Kappa statistics for quantitative evaluation of this tool's validity. Altman et al. suggested a Kappa of 0.8 or more represents substantial agreement, which was achieved in this study (Altman, 1990). There are two main limitations of the unweighted Kappa statistics for validating knowledge. First, it estimates the overall agreement with little or no consideration of the degree or the significance of any individual disagreement. For instance, if there was a persistent disagreement on particular decision rule, the kappa may remain high even though the encountered error is significant. Thus, knowledge error needed to be further stratified according to its significance to ensure the tool as a whole is valid. The other paradox

can be observed when the intra rater agreement kappa is high even though they have disagreed with the decision trees in most occasions. This is due to the fact that kappa is designed to eliminate the agreement by chance rather than discriminating the errors in the classified data. In view of the limitation of the quantitative research model analysis, a qualitative model was introduced and integrated in our study with specific aims and had a higher weight in this validation. The main one was to break down the decision trees into decision rules and assign a level of strength to each rule in the road map of the decision-making. Therefore, this validation is different to previous clinical investigation and essential for validating the knowledge that was used to build the knowledge base.

This was not hypothesis testing study therefore it was not powered and the main focus was on having enough cases to trigger all the rule and qualitatively validate all aspects of the decision rules. However, the quantitative analysis was useful to estimate the approval of the decision rules and set bench mark for the verification. The number of cases included in this study were guided by the results of the clinical investigation and therefore 100 cases were included for the qualitative analysis.

The study had two main limitations. First, human error and subjectivity in rating the cases. We believe this has been limited by having two independent investigators with no familiarity with the tested knowledge and also rating the cases blindly.

Secondly, the included sample not being representative for the population. We minimised that risk by following a rigorous selection of the initial testing set where we only included 35 cases out of 201 and eliminated the rest for being similar. This was supplemented by synthesising 11 cases for the rare clinical presentations. Furthermore, the second set included consecutive cases with no alterations to eliminate the risk of case selection bias. Furthermore, the data were included from only one site and this can

be a limiting factor if the data was not representative. As mentioned above, it has tested all common clinical pathways and only rare presentations had to be synthesised. In the next study, external data and INV for the verification to minimise this limitation.

Prostate cancer is recognised by NICE as a chronic disease and 2014 guidelines suggests offering their follow up in the community. This statement was met with palpable concerns from both the primary and secondary care due to the lack of the required expertise in the community (Watson et al., 2011). A possible solution can be drawn from this study validation methodology, where the integrated knowledge from text guidelines and expert was represented into a valid decision aid. The validation methodology included a pilot trial of the tool on 54 consecutive real cases evaluating its validity and predicts its acceptance by the domain experts.

5.6 Conclusion

This model provided rigorous validation of the expert knowledge with integration of guidelines in the form of decision trees. This is the first model to provide a methodology to quantify the validation of the decision trees and provide a stratification framework to identify the potential errors. The knowledge structure in the form of decision trees can be a platform for developing valid integrated care pathways, guidelines and artificial intelligence based models.

Chapter six - prospective multicentre validation of the prostate cancer follow up expert system

(Chapter published in BJUI Volume122, Issue3 September 2018 Pages 418-426)

6.1 Synopsis

6.1.1 Background

We aimed 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.

6.1.2 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.

6.1.3 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.

6.2 Introduction

In this study we aimed to verify and validate the ES developed in preceding chapter in simulating Urologists in the domain of stable prostate cancer. Therefore, we tested the null hypothesis that the agreement between the ES and the human domain experts is slight to fair (k0 = 0.4) against the alternative: better than moderate (k1 > 0.61). For verification, the ES has to be 100 percent reliable i.e. the same clinical input triggered the system to produce the exact outputs, which was shown above.

6.3 Methods

6.3.1 Description of the prostate cancer system

The ES developed in previous chapters was used as a 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 The system uses principally Microsoft technologies including MS SQL Server database, the Microsoft.NET development framework and the C# programming language along with other web technologies like JavaScript and AJAX.

6.3.2 System security and functional analysis

The expert system is based on a secure remote server held within the NHS 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 in line 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 access their records on the Hospital Information System, including their laboratory test results, before asking the user to fill a problem specific online form. The form includes routine questions enquiring about relevant symptoms as outlined in the NICE guidelines to support the system accurate decision making (NICE, 2014). The system output is in the form of an electronic clinic letter suggesting a tailored treatment plan for the individual patient consultation (appendix 6).

6.3.3 **Study configuration (protocol in appendix 8)**

The CDSS is already part of an existing Clinical Nurse Specialist (CNS) led CaP 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 (appendix 6). This study evaluated an existing service, therefore did not require ethical approval (appendix 3). 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 CNS 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 (CNS 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 CNS if appropriate to do so.

In this study, we only aimed to evaluated the system 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 (appendix 6):

1. History of disease and treatment.
2. Current symptoms and lab test results.
3. System analysis of clinical data.
4. System suggested outcome.
5. Nurse Note's 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 CNS free text entry of their own clinical assessment and plan (component 5 in appendix 6)

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. not CaP etc.).

6.3.3.1 Eligibility criteria

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

6.3.3.2 Inclusion criteria

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

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

With or without (optional requirement):

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

6.3.3.3 Exclusion criteria

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

In order to reach both end points, the study divided into two branches after case selection (Figure 6.1). In the primary end point pathway, the included cases were reviewed by two independent clinical investigators after eliminating the system analysis and suggestions (Components 3 and 4 samples in appendix 6). Each investigator independently assigned an outcome for each case (appendix 7) according to the disease history, current results and symptoms blinded from the system recommendation. Both investigators are consultants with different interest and practicing in separate regions.

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) 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 retest by a new sample of cases to re-estimate reliability until 100% precision was obtained (secondary end point).

6.3.4 **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 ($k0 = 0.4$) against the alternative that it was better than moderate ($k1 > 0.61$) (Cohen, 1960). Kappa was estimated between each investigator (INV1 & INV2) and then the CDSS between both investigators.

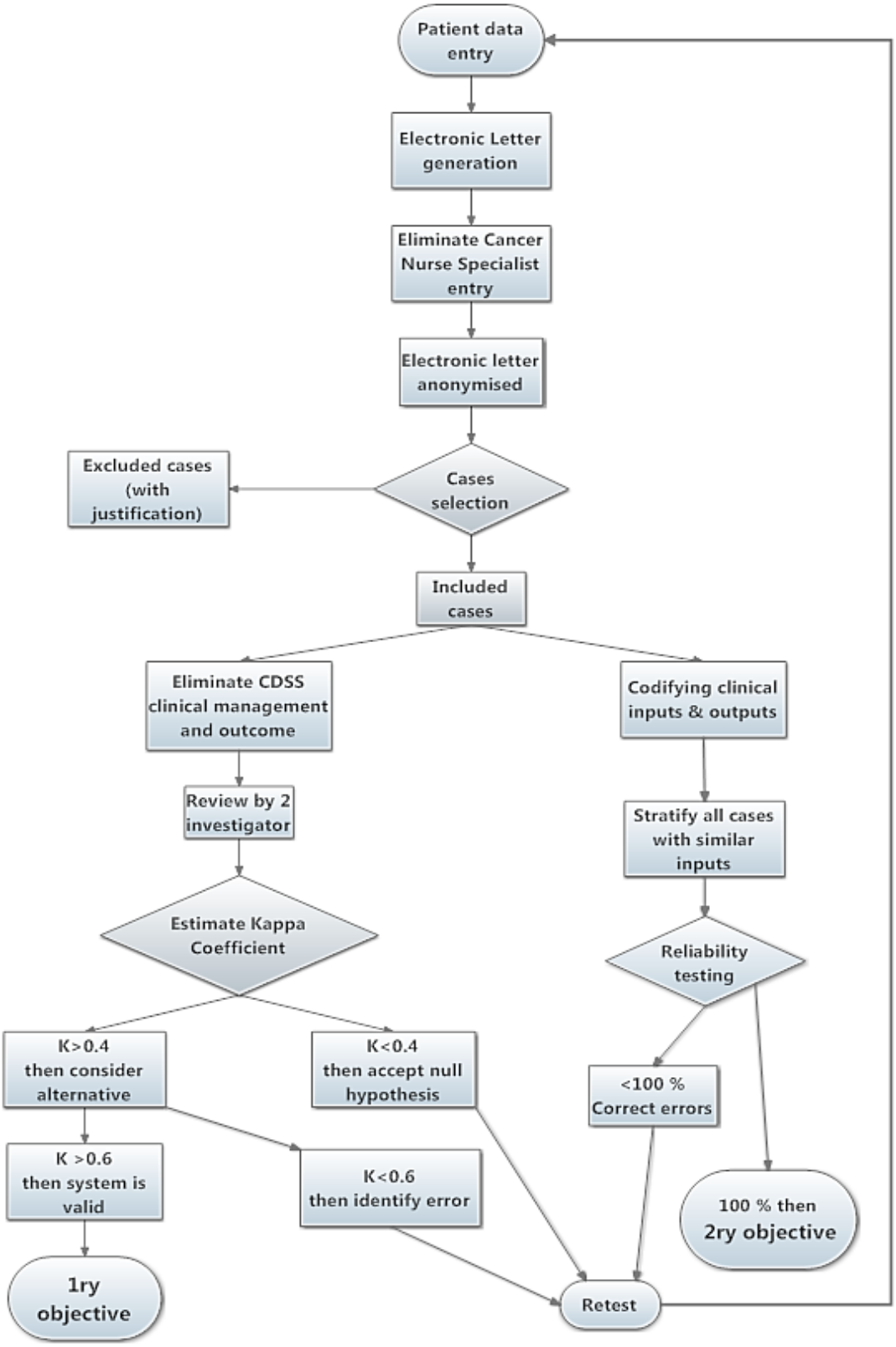


Figure 6.1 Live validation study flow chart

6.4 **Results**

The study included data of 200 patients seen in either hospitals (100 each Figure 6.2). All patients had CaP and mean and median age was 75 (range 51 to 94) with a mean and mode Gleason score of 7. The risk stratifications according to D'Amico's classification (D'Amico et al., 1998) of their disease on diagnosis demonstrated 96 (48%) patients in the high risk group and 78 (39%) & 26 (13%) in the intermediate and low risk group respectively (Figure 6.3Figure 5.2). Radiotherapy had been used to treat 128 cases out of the total (Figure 6.4).

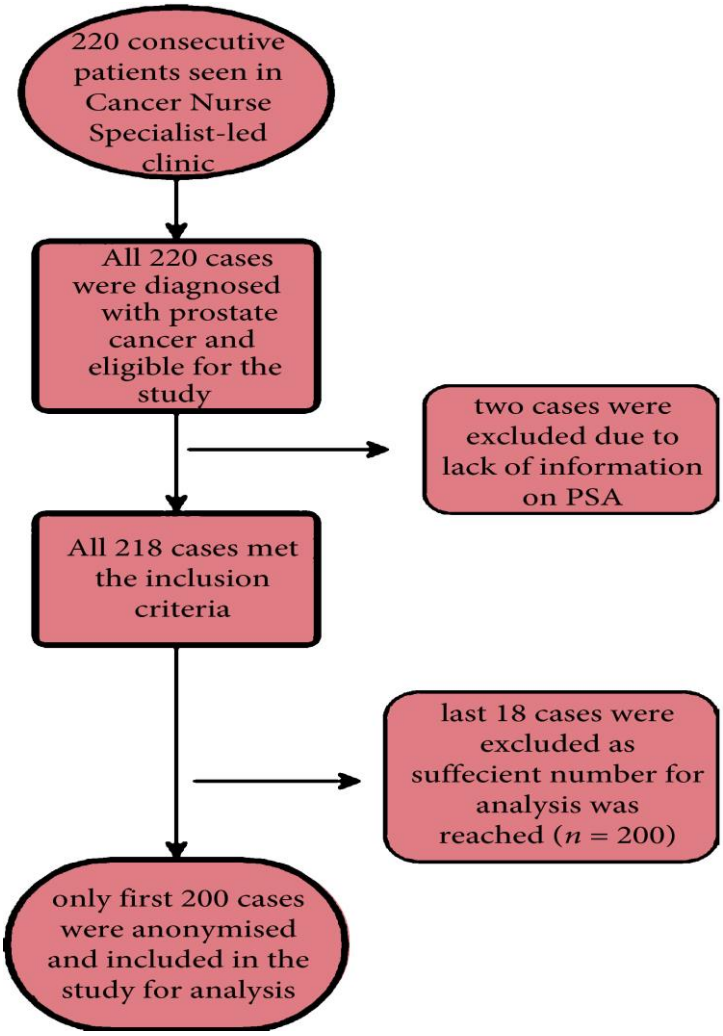


Figure 6.2 Included cases flow chart.

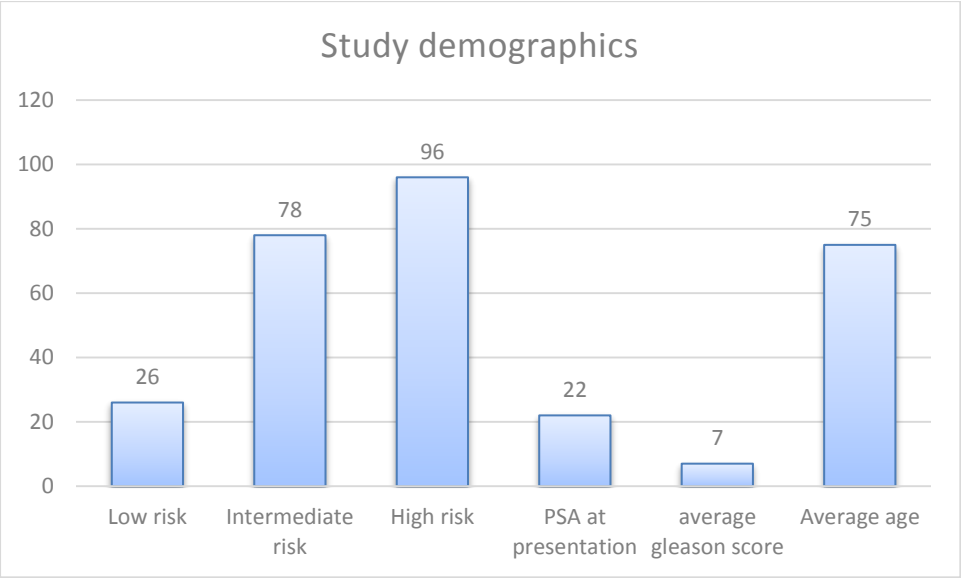


Figure 6.3 included cases demographics
Cases were stratified according to their data on the initial diagnosis of CaP following D'Amico's classification criteria. Most were in the intermediate and the high risk group.

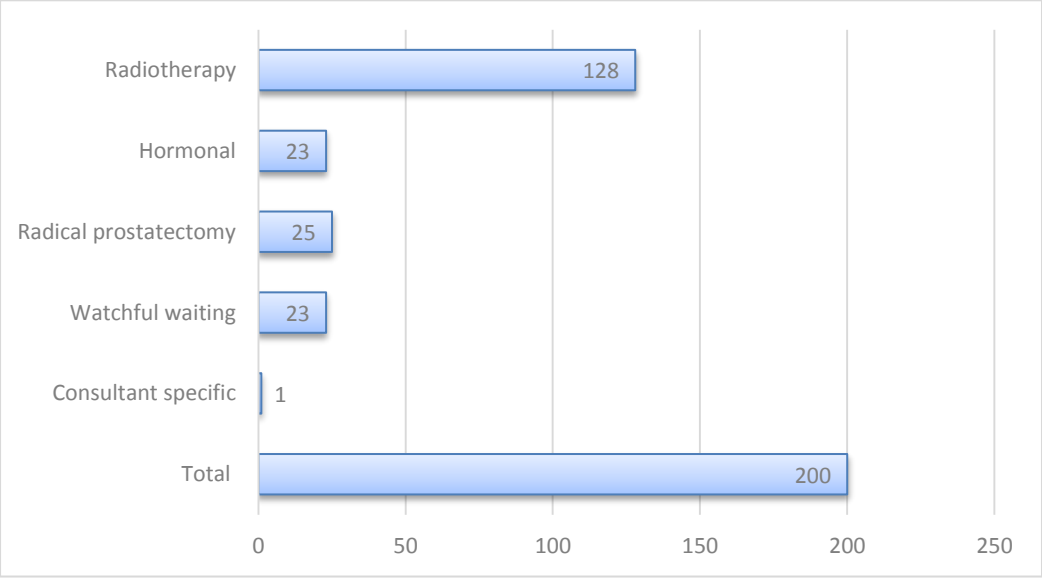


Figure 6.4 Cases stratification according to their treatment
The stratification was according to the most significant treatment on which we calculate the biochemical failure and disease progression. Radiotherapy was the most common as this group usually requires longer follow up than patient post-surgical treatment

Kappa statistics were estimated to test the hypothesis. The unweighted kappa for INV1 INV2 was 0.8, and with quadratic weighting the kappa remained, demonstrating

substantial agreement between CDSS and investigator at 0.86 and 0.96 respectively for INV1 and INV2 respectively. This refuted the null hypothesis as k0 > 0.4 and alternative hypothesis k1 > 0.6 was accepted with the conclusion that the CDSS is validation and its assigned outcomes are acceptable by the domain experts (Table 6.1,Table 6.2 &Table 6.3).

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 6.1 Kappa statistics for investigator 1 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*

Table 6.2 Kappa statistics for investigator 2 and CDSS

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

Table 6.3 Kappa statistics for investigator 1 and investigator 2

Both investigators identified 31 cases with possible disease progression or recurrence (INV1 = 25, INV2 = 29). Most of those cases were biochemical failure (BCF) post radiotherapy treatment (21 out of the total of 31) based on either ASTRO (Bellera et al., 2009, Kuban et al., 2006, Horwitz et al., 2005) or PHOENIX (ASTRO, 1997) criteria or both. Six cases were on watchful waiting had either high or rising PSA and were identified as disease progression (those patients did not receive any treatment). Two

patients had radical surgery with detectable PSA and one was treated with hormone ablation, with significant PSA doubling time. CDSS identified all of the above cases as disease progression or recurrence except one case who was only classified by INV1 as potential recurrence post radiotherapy. They had only two consecutive rises in their PSA, thus not identified as BCF by CDSS or INV2.

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

6.5 Discussion

This 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. This is the first study 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 health care because of lack of confidence in these systems (Ammenwerth et al., 2013). Standalone software as CDSS are recognised by European regulatory bodies as medical devices and they require registration via strict criteria (93/42/EEC, 2007). These criteria are outlined in the declarations of the International Organization for Standardization (ISO) 13485 (quality management systems), 14971 (risk management) and 62304 (software development cycle). The latter addresses the standards for testing that should met in the validation. In this study, we quantified the validity by estimating the kappa statistics in a 2 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 (93/42/EEC, 2007).

It has been estimated that CaP 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 (Luengo-Fernandez et al., 2013). 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 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 (Watson et al., 2011). The system developers believe that the valid system can be used by any healthcare worker regardless 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 health care, 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 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 CNS follow up clinic by the MDT and may explain the low recurrence rate in this 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 CaP 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 has been 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 (Turing, 1950, Okeefe and Oleary, 1993). Kappa statistics is more appropriate for this type of validation as they estimate the agreement between human and machine cognitive function and has been applied by other studies to validate expert systems (Lopes et al., 2013, Torshizi et al., 2014). The sensitivity and specificity is only accurate if compared the gold standard as 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 (Pandey and Mishra, 2009).

The long term outcome for those cases was not examined as part of this 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.

6.6 Conclusion

The CDSS demonstrated high validity and accuracy in its decision making. This tool has potential for safe use in supporting follow up of surgical oncology in the primary and secondary care. Mixed methodology approach is required to perform the mandatory system validation and verification.

Thesis Conclusion

7.1 Summary

Artificial intelligence was built by scientists to replicate human knowledge in order to simulate human intelligence and solve complex but mundane problems. In industry, they have developed over the years to become a standard in several different fields (Liao, 2005). Their development and uptake lagged in health care, and despite leading by the first published rule based ES "MYCIN (Shortliffe et al., 1975).

Despite their plentiful supply, as demonstrated in our systematic review (chapter two), there has been an a huge gap between production and uptake in healthcare (Ammenwerth et al., 2013). Reasons suggested for lack of uptake include a lack of formality in their development and a lack of validation are leading causes (Ammenwerth et al., 2013) as well as a lack of system evaluation. A valid and accurate system that is not seen as valuable and reliable by the user will remain unpopular, therefore system evaluation is a crucial step but found to be defective in almost all of the ES in the review above.

A second observation from our systematic review was the lack of system verification which is a critical for performing the conformity check and registering the system with the regulatory bodies (93/42/EEC, 2007).

The lack of transparency of some model's calculations, as in ANN, results in the black box phenomenon and consequent lack of trust in the outputs.

Another novel approach was applied to validate the knowledge before finalising the knowledge base (in chapter five). In this study the knowledge was highly refined while

being structure in a hand written rule based format. This has simplified further validation and verification by deciding the statistical test and estimating the hypothesis testing. The study also quantified the knowledge validation and allowed further error identification and correction.

Further challenge in these system development is system integration with the NHS networks and MHRA approval. The MRHA regulation for software registration varies according to its intended use and application. As for our CDSS, its application in the hospital clinic as intelligent database and electronic letter generator does not need registration since its calculations were being supervised by a CNS. However, for the purpose of being standalone software where its suggestions and calculations are unsupervised, the system will need to be registered as category 1a medical device (in vitro device, conformity check required) (93/42/EEC, 2007). Thus, the system was applied in two units to act as an intelligent database under close supervision from CNS then real time data were collected for live testing. The data acquired from these clinic were highly useful for performing the validation and verification testing that satisfies MHRA regulations.

7.2 Practice implication

The systematic review demonstrated the ES versatility and diversity of applications. One of their advanced application is research variables prediction for enhancing the statistical regression analysis for research domain with noisy variables.

Expert systems application can improve decision making specially in areas of high demand as prostate cancer follow up. Their application can have improve clinical outcome and patients' care as demonstrated by systematic reviews of clinical decision support systems (Garg et al., 2005, Kawamoto et al., 2005). They have an added value

in electronic record generation and keeping. They have the ability to retain and process large sets of data which can enhance clinical auditing, research and patients' care.

7.3 Limitation

These types of studies are always limited by the quality of knowledge acquired from the domain experts developing the system then the experts performing the hypothesis testing. Other human factor shortcomings, include bias and subjectivity.

The human expert judgment was used as gold standard in our studies due to the limitation of using other tests as MRI and TRUS biopsies. It would be unethical to subject patients to such invasive and expensive tests where they are not routinely indicated in follow up. The alternative methodology that we applied is not uncommon and has been reported by Turing for testing machine intelligence (Turing, 1950). Furthermore, the time frame of system development is also a limiting factors as the knowledge is expected to change. Thus it is unrealistic to develop a system in a RCT which would take years to perform and then the end result of the system be invalid due to the change of scientific evidence and knowledge.

7.4 Future

In general, further research should be undertaken in system evaluation to determine the system usability before its development as in other industrial systems. This would allow healthcare system developers to identify domains with high demands for CDSS.

More research is required in medical knowledge acquisition and further validation as this is considered the bottleneck for these systems development.

Furthermore, even though we have proved that our CDSS is valid and can produce treatment plans comparable to domain experts in blinded studies we have not proven

that this service is more cost effective in comparison to current secondary care service. Helgesen (2000) and Faithful (2001) evaluated their nurse led follow up of post radiotherapy patients and their health economic evaluation found their service to be cost effective and similar types of studies should be performed with our system as the subject (Faithfull et al., 2001) (Helgesen et al., 2000).

ES registration needs further clarification from the regulatory bodies to enhance their development and application in health care. Also more formality of testing is required and education on the International Organization for Standardization (ISO) requirement for health care systems developers to follow.

7.5 Conclusion

Artificial intelligence have proven to be highly versatile and accurate tools for supporting decision making and complex problem solving. They have a complex development and validation which is challenged by lack of formality. The methods reported in this thesis allowed the development and testing of CDSS in CaP to the standards of the conformity testing. The CDSS can support health care workers to provide safe care for patients with CaP in the community.

References

93/42/EEC. 2007. *RE: DIRECTIVE 2007/47/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 5 September 2007.*

ABBOD, M. F., CATTO, J. W. F., CHEN, M., LINKENS, D. A. & HAMDY, F. C. 2004. ARTIFICIAL INTELLIGENCE FOR THE PREDICTION OF BLADDER CANCER. *Biomedical Engineering: Applications, Basis and Communications,* 16**,** 49-58.

ABBOD, M. F., CATTO, J. W. F., LINKENS, D. A., WILD, P. J., HERR, A., WISSMANN, C., PILARSKY, C., HARTMANN, A. & HAMDY, F. C. Artificial Intelligence Technique for Gene Expression Profiling of Urinary Bladder Cancer. Intelligent Systems, 2006 3rd International IEEE Conference on, Sept. 2006 2006. 646-651.

ACHOUR, S. L., DOJAT, M., RIEUX, C., BIERLING, P. & LEPAGE, E. 2001. Application of information technology - A UMLS-based knowledge acquisition tool for rule-based clinical decision support system development. *Journal of the American Medical Informatics Association,* 8**,** 351-360.

ACKOFF, R. L. 1989. "From Data to Wisdom". *Journal of Applied Systems Analysis,* Volume 16**,** 3-9.

ADRION, W. R., BRANSTAD, M. A. & CHERNIAVSKY, J. C. 1982. VALIDATION, VERIFICATION, AND TESTING OF COMPUTER SOFTWARE. *Computing Surveys,* 14**,** 159-192.

AHMED, H. U., EL-SHATER BOSAILY, A., BROWN, L. C., GABE, R., KAPLAN, R., PARMAR, M. K., COLLACO-MORAES, Y., WARD, K., HINDLEY, R. G., FREEMAN, A., KIRKHAM, A. P., OLDROYD, R., PARKER, C. & EMBERTON, M. 2017. Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study. *The Lancet,* 389**,** 815-822.

ALBERTSEN, P. C. 2015. Observational studies and the natural history of screen-detected prostate cancer. *Curr Opin Urol,* 25**,** 232-7.

ALTMAN, D. G. 1990. *Practical statistics for medical research*, CRC Press.

ALTUNAY, S., TELATAR, Z., EROGUL, O. & AYDUR, E. 2009. A new approach to urinary system dynamics problems: Evaluation and classification of uroflowmeter signals using artificial neural networks. *Expert Systems with Applications,* 36**,** 4891-4895.

AMMENWERTH, E., NYKANEN, P., RIGBY, M. & DE KEIZER, N. 2013. Clinical decision support systems: Need for evidence, need for evaluation. *Artificial Intelligence in Medicine,* 59**,** 1-3.

ANAGNOSTOU, T., REMZI, M., LYKOURINAS, M. & DJAVAN, D. 2003. Artificial neural networks for decision-making in urologic oncology. *European Urology,* 43**,** 596-603.

ARLEN, A. M., ALEXANDER, S. E., WALD, M. & COOPER, C. S. Computer model predicting breakthrough febrile urinary tract infection in children with primary vesicoureteral reflux. *Journal of Pediatric Urology*.

ASHRAFI, N., KUILBOER, J. P. & WAGNER, J. M. 1995. EXPERT-SYSTEMS RELIABILITY - A LIFE-CYCLE APPROACH. *Information & Management,* 28**,** 405-414.

ASTRO 1997. Consensus statement: guidelines for PSA following radiation therapy. American Society for Therapeutic Radiology and Oncology Consensus Panel. *International journal of radiation oncology, biology, physics,* 37**,** 1035-41.

BABAIAN, R. J., FRITSCHE, H., AYALA, A., BHADKAMKAR, V., JOHNSTON, D. A., NACCARATO, W. & ZHANG, Z. 2000. Performance of a neural network in detecting prostate cancer in the prostate-specific antigen reflex range of 2.5 to 4.0 ng/ml. *Urology,* 56**,** 1000-1006.

BABAIAN, R. J., FRITSCHE, H. A., ZHANG, Z., ZHANG, K. H., MADYASTHA, K. R. & BARNHILL, S. D. 1998. Evaluation of ProstAsure index in the detection of prostate cancer: A preliminary report. *Urology,* 51**,** 132-136.

BAGLI, D. J., AGARWAL, S. K., VENKATESWARAN, S., SHUCKETT, B., KHOURY, A. E., MERGUERIAN, P. A., MCLORIE, G. A., LIU, K. & NIEDERBERGER, C. S. 1998. Artificial neural networks in pediatric urology: Prediction of sonographic outcome following pyeloplasty. *Journal of Urology,* 160**,** 980-983.

BAGSHAW, M. A., KAPLAN, H. S. & SAGERMAN, R. H. 1965. LINEAR ACCELERATOR SUPERVOLTAGE RADIOTHERAPY .7. CARCINOMA OF PROSTATE. *Radiology,* 85**,** 121-&.

BAGSHAW, M. A., RAY, G. R., PISTENMA, D. A., CASTELLINO, R. A. & MEARES, E. M. 1975. EXTERNAL BEAM RADIATION-THERAPY OF PRIMARY CARCINOMA OF PROSTATE. *Cancer,* 36**,** 723-728.

BANCROFT, E. K., PAGE, E. C., CASTRO, E., LILJA, H., VICKERS, A., SJOBERG, D., ASSEL, M., FOSTER, C. S., MITCHELL, G., DREW, K., MÆHLE, L., AXCRONA, K., EVANS, D. G., BULMAN, B., ECCLES, D., MCBRIDE, D., VAN ASPEREN, C., VASEN, H., KIEMENEY, L. A., RINGELBERG, J., CYBULSKI, C., WOKOLORCZYK, D., SELKIRK, C., HULICK, P. J., BOJESEN, A., SKYTTE, A.-B., LAM, J., TAYLOR, L., OLDENBURG, R., CREMERS, R., VERHAEGH, G., VAN ZELST-STAMS, W. A., OOSTERWIJK, J. C., BLANCO, I., SALINAS, M., COOK, J., ROSARIO, D. J., BUYS, S., CONNER, T., AUSEMS, M. G., ONG, K.-R., HOFFMAN, J., DOMCHEK, S., POWERS, J., TEIXEIRA, M. R., MAIA, S., FOULKES, W. D., TAHERIAN, N., RUIJS, M., DEN ENDEN, A. T. H.-V., IZATT, L., DAVIDSON, R., ADANK, M. A., WALKER, L., SCHMUTZLER, R., TUCKER, K., KIRK, J., HODGSON, S., HARRIS, M., DOUGLAS, F., LINDEMAN, G. J., ZGAJNAR, J., TISCHKOWITZ, M., CLOWES, V. E., SUSMAN, R., RAMÓN Y CAJAL, T., PATCHER, N., GADEA, N., SPIGELMAN, A., VAN OS, T., LILJEGREN, A., SIDE, L., BREWER, C., BRADY, A. F., DONALDSON, A., STEFANSDOTTIR, V., FRIEDMAN, E., CHEN-SHTOYERMAN, R., AMOR, D. J., COPAKOVA, L., BARWELL, J., GIRI, V. N., MURTHY, V., NICOLAI, N., TEO, S.-H., GREENHALGH, L., STROM, S., HENDERSON, A., MCGRATH, J., GALLAGHER, D., AARONSON, N., ARDERN-JONES, A., BANGMA, C., DEARNALEY, D., COSTELLO, P., EYFJORD, J., ROTHWELL, J., FALCONER, A., GRONBERG, H., HAMDY, F. C., et al. 2014. Targeted Prostate Cancer Screening in BRCA1 and BRCA2 Mutation Carriers: Results from the Initial Screening Round of the IMPACT Study. *European Urology,* 66**,** 489-499.

BASSI, P., SACCO, E., DE MARCO, V., ARAGONA, M. & VOLPE, A. 2007. Prognostic accuracy of an artificial neural network in patients undergoing radical cystectomy for bladder cancer: a comparison with logistic regression analysis. *Bju International,* 99**,** 1007-1012.

BATUELLO, J. T., GAMITO, E. J., CRAWFORD, E. D., HAN, M., PARTIN, A. W., MCLEOD, D. G. & O'DONNELL, C. 2001. Artificial neural network model for the assessment of lymph node spread in patients with clinically localized prostate cancer. *Urology,* 57**,** 481-485.

BAUS 2017. Analyses of Radical Prostatectomies performed between

January 1st and December 31st 2016.

BELIGIANNIS, G., HATZILYGEROUDIS, I., KOUTSOJANNIS, C. & PRENTZAS, J. 2006. A GA driven intelligent system for medical diagnosis. *Knowledge-Based Intelligent Information and Engineering Systems, Pt 1, Proceedings,* 4251**,** 968-975.

BELLERA, C. A., HANLEY, J. A., JOSEPH, L. & ALBERTSEN, P. C. 2009. A STATISTICAL EVALUATION OF RULES FOR BIOCHEMICAL FAILURE AFTER RADIOTHERAPY IN MEN TREATED FOR PROSTATE CANCER. *International Journal of Radiation Oncology Biology Physics,* 75**,** 1357-1363.

BENBASAT, I. & DHALIWAL, J. S. 1989. A framework for the validation of knowledge acquisition. *Knowledge Acquisition,* 1**,** 215-233.

BENECCHI, L. 2006. Neuro-fuzzy system for prostate cancer diagnosis. *Urology,* 68**,** 357-361.

BILL-AXELSON, A., HOLMBERG, L., GARMO, H., TAARI, K., BUSCH, C., NORDLING, S., HAGGMAN, M., ANDERSSON, S. O., ANDREN, O., STEINECK, G., ADAMI, H. O. & JOHANSSON, J. E.

2018. Radical Prostatectomy or Watchful Waiting in Prostate Cancer - 29-Year Follow-up. *N Engl J Med,* 379**,** 2319-2329.

BILL-AXELSON, A., HOLMBERG, L., RUUTU, M., GARMO, H., STARK, J. R., BUSCH, C., NORDLING, S., HAGGMAN, M., ANDERSSON, S. O., BRATELL, S., SPANGBERG, A., PALMGREN, J., STEINECK, G., ADAMI, H. O., JOHANSSON, J. E. & INVESTIGATORS, S.-. 2011. Radical Prostatectomy versus Watchful Waiting in Early Prostate Cancer. *New England Journal of Medicine,* 364**,** 1708-1717.

BINDER, J. & KRAMER, W. 2001. Robotically-assisted laparoscopic radical prostatectomy. *BJU Int,* 87**,** 408-10.

BINIK, Y. M., SERVANSCHREIBER, D., FREIWALD, S. & HALL, K. S. K. 1988. INTELLIGENT COMPUTER-BASED ASSESSMENT AND PSYCHOTHERAPY - AN EXPERT SYSTEM FOR SEXUAL DYSFUNCTION. *Journal of Nervous and Mental Disease,* 176**,** 387-400.

BLACKARD, C. E. 1975. The Veterans' Administration Cooperative Urological Research Group studies of carcinoma of the prostate: a review. *Cancer Chemother Rep,* 59**,** 225-7.

BOEGL, K., ADLASSNIG, K. P., HAYASHI, Y., ROTHENFLUH, T. E. & LEITICH, H. 2004. Knowledge acquisition in the fuzzy knowledge representation framework of a medical consultation system. *Artificial Intelligence in Medicine,* 30**,** 1-26.

BOLLA, M., MAINGON, P., CARRIE, C., VILLA, S., KITSIOS, P., POORTMANS, P. M., SUNDAR, S., VAN DER STEEN-BANASIK, E. M., ARMSTRONG, J., BOSSET, J. F., HERRERA, F. G., PIETERS, B., SLOT, A., BAHL, A., BEN-YOSEF, R., BOEHMER, D., SCRASE, C., RENARD, L., SHASH, E., COENS, C., VAN DEN BERGH, A. C. & COLLETTE, L. 2016. Short Androgen Suppression and Radiation Dose Escalation for Intermediate- and High-Risk Localized Prostate Cancer: Results of EORTC Trial 22991. *J Clin Oncol,* 34**,** 1748-56.

BORQUE, A., SANZ, G., ALLEPUZ, C., PLAZA, L., GIL, P. & RIOJA, L. A. 2001. The use of neural networks and logistic regression analysis for predicting pathological stage in men undergoing radical prostatectomy: a population based study. *J Urol,* 166**,** 1672-8.

BOTOCA, C., BARDAN, R., BOTOCA, M. & ALEXA, F. 2009. Prediction of Prostate Capsule Penetration using Neural Networks. *Proceedings of the 8th Wseas International Conference on Computational Intelligence, Man-Machine Systems and Cybernetics (Cimmacs '09)***,** 108-111.

BOYINGTON, A. R., WILDEMUTH, B. M., DOUGHERTY, M. C. & HALL, E. P. 2004. Development of a computer-based system for continence health promotion. *Nursing Outlook,* 52**,** 241-247.

BRAY, F., LORTET-TIEULENT, J., FERLAY, J., FORMAN, D. & AUVINEN, A. 2010. Prostate cancer incidence and mortality trends in 37 European countries: An overview. *European Journal of Cancer,* 46**,** 3040-3052.

BREAST CANC LINKAGE, C. 1999. Cancer risks in BRCA2 mutation carriers. *Journal of the National Cancer Institute,* 91**,** 1310-1316.

BREWSTER, D. H., FRASER, L. A., HARRIS, V. & BLACK, R. J. 2000. Rising incidence of prostate cancer in Scotland: increased risk or increased detection? *Bju International,* 85**,** 463-472.

BRIGANTI, A., FOSSATI, N., CATTO, J. W. F., CORNFORD, P., MONTORSI, F., MOTTET, N., WIRTH, M. & VAN POPPEL, H. 2018. Active Surveillance for Low-risk Prostate Cancer: The European Association of Urology Position in 2018. *European Urology,* 74**,** 357-368.

BRUNER, D. W., MOORE, D., PARLANTI, A., DORGAN, J. & ENGSTROM, P. 2003. Relative risk of prostate cancer for men with affected relatives: Systematic review and meta-analysis. *International Journal of Cancer,* 107**,** 797-803.

BUCHNER, A., KENDLBACHER, M., NUHN, P., TULLMANN, C., HASEKE, N., STIEF, C. G. & STAEHLER, M. 2012. Outcome Assessment of Patients With Metastatic Renal Cell Carcinoma Under Systemic Therapy Using Artificial Neural Networks. *Clinical Genitourinary Cancer,* 10**,** 37-42.

BUCHNER, A., MAY, M., BURGER, M., BOLENZ, C., HERRMANN, E., FRITSCHE, H. M., ELLINGER, J., HOFNER, T., NUHN, P., GRATZKE, C., BROOKMAN-MAY, S., MELCHIOR, S., PETER, J., MORITZ, R., TILKI, D., GILFRICH, C., ROIGAS, J., ZACHARIAS, M., HOHENFELLNER, M., HAFERKAMP, A., TROJAN, L., WIELAND, W. F., MULLER, S. C., STIEF, C. G. & BASTIAN, P. J. 2013. Prediction of outcome in patients with urothelial carcinoma of the bladder following radical cystectomy using artificial neural networks. *Ejso,* 39**,** 372-379.

BYRD, T. A., COSSICK, K. L. & ZMUD, R. W. 1992. A SYNTHESIS OF RESEARCH ON REQUIREMENTS ANALYSIS AND KNOWLEDGE ACQUISITION TECHNIQUES. *Mis Quarterly,* 16**,** 117-138.

CAI, T., CONTI, G., LORENZINI, M. & BARTOLETTI, R. 2007a. Artificial intelligences in urological practice: the key to success? *Annals of Oncology,* 18**,** 604-U10.

CAI, T., CONTI, G., NESI, G., LORENZINI, M., MONDAINI, N. & BARTOLETTI, R. 2007b. Artificial intelligence for predicting recurrence-free probability of non-invasive high-grade urothelial bladder cell carcinoma. *Oncol Rep,* 18**,** 959-64.

CAI, T., NESI, G., TINACCI, G., GIUBILEI, G., GAVAZZI, A., MONDAINI, N., ZINI, E. & BARTOLETTI, R. 2011. Clinical importance of lymph node density in predicting outcome of prostate cancer patients. *J Surg Res,* 167**,** 267-72.

CANCERRESEARCHUK. 2016a. *Prostate cancer incidence statistics* [Online]. Available: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/prostate-cancer/incidence#heading-Two.

CANCERRESEARCHUK. 2016b. *Prostate cancer mortality* [Online]. Available: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/prostate-cancer#heading-One.

CASTANHO, M. J. P., HERNANDES, F., DE RE, A. M., RAUTENBERG, S. & BILLIS, A. 2013. Fuzzy expert system for predicting pathological stage of prostate cancer. *Expert Systems with Applications,* 40**,** 466-470.

CATTO, J. W. F., ABBOD, M. F., LINKENS, D. A. & HAMDY, F. C. 2006. Neuro-fuzzy modeling: An accurate and interpretable method for predicting bladder cancer progression. *Journal of Urology,* 175**,** 474-479.

CATTO, J. W. F., ABBOD, M. F., LINKENS, D. A., LARRE, S., ROSARIO, D. J. & HAMDY, F. C. 2009. Neurofuzzy Modeling to Determine Recurrence Risk Following Radical Cystectomy for Nonmetastatic Urothelial Carcinoma of the Bladder. *Clinical Cancer Research,* 15**,** 3150-3155.

CATTO, J. W. F., ABBOD, M. F., WILD, P. J., LINKENS, D. A., PILARSKY, C., REHMAN, I., ROSARIO, D. J., DENZINGER, S., BURGER, M., STOEHR, R., KNUECHEL, R., HARTMANN, A. & HAMDY, F. C. 2010. The Application of Artificial Intelligence to Microarray Data: Identification of a Novel Gene Signature to Identify Bladder Cancer Progression. *European Urology,* 57**,** 398-406.

CATTO, J. W. F., LINKENS, D. A., ABBOD, M. F., CHEN, M. Y., BURTON, J. L., FEELEY, K. M. & HAMDY, F. C. 2003. Artificial intelligence in predicting bladder cancer outcome: A comparison of neuro-fuzzy modeling and artificial neural networks. *Clinical Cancer Research,* 9**,** 4172-4177.

CHANG, C. L., COMBS, J. B. & STACHOWITZ, R. A. 1990. A report on the Expert Systems Validation Associate (EVA). *Expert Systems with Applications,* 1**,** 217-230.

CHANG, P. L., LI, Y. C., WANG, T. M., HUANG, S. T., HSIEH, M. L. & TSUI, K. H. 1999. Evaluation of a decision-support system for preoperative staging of prostate cancer. *Medical Decision Making,* 19**,** 419-427.

CHIU, J. S., WANG, Y. F., SU, Y. C., WEI, L. H., LIAO, J. G. & LI, Y. C. 2009. Artificial Neural Network to Predict Skeletal Metastasis in Patients with Prostate Cancer. *Journal of Medical Systems,* 33**,** 91-100.

CINAR, M., ENGIN, M., ENGIN, E. Z. & ATESCI, Y. Z. 2009. Early prostate cancer diagnosis by using artificial neural networks and support vector machines. *Expert Systems with Applications,* 36**,** 6357-6361.

CLOSE, A., ROBERTSON, C., RUSHTON, S., SHIRLEY, M., VALE, L., RAMSAY, C. & PICKARD, R. 2013. Comparative Cost-effectiveness of Robot-assisted and Standard Laparoscopic Prostatectomy as Alternatives to Open Radical Prostatectomy for Treatment of Men with Localised Prostate Cancer: A Health Technology Assessment from the Perspective of the UK National Health Service. *European Urology,* 64**,** 361-369.

COHEN, J. 1960. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement,* 20**,** 37-46.

COHEN, J. 1968. Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological Bulletin,* 70**,** 213-220.

COLLIN, S. M., MARTIN, R. M., METCALFE, C., GUNNELL, D., CALBERTSEN, P., NEAL, D., HAMDY, F., STEPHENS, P., LANE, J. A., MOORE, R. & DONOVAN, J. 2008. Prostate-cancer mortality in the USA and UK in 1975-2004: an ecological study. *Lancet Oncology,* 9**,** 445-452.

COOKSON, M. S., AUS, G., BURNETT, A. L., CANBY-HAGINO, E. D., D'AMICO, A. V., DMOCHOWSKI, R. R., ETON, D. T., FORMAN, J. D., GOLDENBERG, S. L., HERNANDEZ, J., HIGANO, C. S., KRAUS, S. R., MOUL, J. W., TANGEN, C., THRASHER, J. B. & THOMPSON, I. 2007. Variation in the definition of biochemical recurrence in patients treated for localized prostate cancer: The American Urological Association Prostate Guidelines for Localized Prostate Cancer update panel report and recommendations for a standard in the reporting of surgical outcomes. *Journal of Urology,* 177**,** 540-545.

COSMA, G., ACAMPORA, G., BROWN, D., REES, R. C., KHAN, M. & POCKLEY, A. G. 2016. Prediction of Pathological Stage in Patients with Prostate Cancer: A Neuro-Fuzzy Model. *PLoS ONE,* 11**,** e0155856.

CUMMINGS, J. M., BOULLIER, J. A., IZENBERG, S. D., KITCHENS, D. M. & KOTHANDAPANI, R. V. 2000. Prediction of spontaneous ureteral calculous passage by an artificial neural network. *Journal of Urology,* 164**,** 326-328.

D'AMICO, A. V., MOUL, J., CARROLL, P. R., SUN, L., LUBECK, D. & CHEN, M.-H. 2004. Prostate specific antigen doubling time as a surrogate end point for prostate cancer specific mortality following radical prostatectomy or radiation therapy. *The Journal of urology,* 172**,** S42-6; discussion S46-7.

D'AMICO, A. V., WHITTINGTON, R., MALKOWICZ, S. B., SCHULTZ, D., BLANK, K., BRODERICK, G. A., TOMASZEWSKI, J. E., RENSHAW, A. A., KAPLAN, I., BEARD, C. J. & WEIN, A. 1998. Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. *JAMA,* 280**,** 969-74.

DAL MORO, F., ABATE, A., LANCKRIET, G. R. G., ARANDJELOVIC, G., GASPARELLA, P., BASSI, P., MANCINI, M. & PAGANO, F. 2006. A novel approach for accurate prediction of spontaneous passage of ureteral stones: Support vector machines. *Kidney International,* 69**,** 157-160.

DALESIO, O., VAN TINTEREN, H., CLARKE, M., PETO, R., SCHRODER, F. H., DECHERING, I., EVANS, V., GODWIN, J., BLUMENSTEIN, B. A., CRAWFORD, E. D., DENIS, L., HALL, R., HILL, C., IVERSEN, P., SHIPLEY, W. U., SOLOWAY, M., SYLVESTER, R. & PROSTATE CANC TRIALISTS, C. 2000. Maximum androgen blockade in advanced prostate cancer: an overview of the randomised trials. *Lancet,* 355**,** 1491-1498.

DALL'ERA, M. A., ALBERTSEN, P. C., BANGMA, C., CARROLL, P. R., CARTER, H. B., COOPERBERG, M. R., FREEDLAND, S. J., KLOTZ, L. H., PARKER, C. & SOLOWAY, M. S. 2012. Active Surveillance for Prostate Cancer: A Systematic Review of the Literature. *European Urology,* 62**,** 976-983.

DAVIS, R. & BUCHANAN, B. C. 1977. Meta-level knowledge: overview and applications. *Proceedings of the 5th international joint conference on Artificial intelligence - Volume 2.* Cambridge, USA: Morgan Kaufmann Publishers Inc.

DE BONO, J. S., LOGOTHETIS, C. J., MOLINA, A., FIZAZI, K., NORTH, S., CHU, L., CHI, K. N., JONES, R. J., GOODMAN, O. B., SAAD, F., STAFFURTH, J. N., MAINWARING, P., HARLAND, S., FLAIG, T. W., HUTSON, T. E., CHENG, T., PATTERSON, H., HAINSWORTH, J. D., RYAN, C. J., STERNBERG, C. N., ELLARD, S. L., FLÉCHON, A., SALEH, M., SCHOLZ, M., EFSTATHIOU, E., ZIVI, A., BIANCHINI, D., LORIOT, Y., CHIEFFO, N., KHEOH, T., HAQQ, C. M. & SCHER, H. I. 2011. Abiraterone and Increased Survival in Metastatic Prostate Cancer. *New England Journal of Medicine,* 364**,** 1995-2005.

DE PAULA CASTANHO, M. J., DE BARROS, L. C., YAMAKAMI, A. & VENDITE, L. L. 2008. Fuzzy expert system: An example in prostate cancer. *Applied Mathematics and Computation,* 202**,** 78-85.

DEARNALEY, D., SYNDIKUS, I., MOSSOP, H., KHOO, V., BIRTLE, A., BLOOMFIELD, D., GRAHAM, J., KIRKBRIDE, P., LOGUE, J., MALIK, Z., MONEY-KYRLE, J., O'SULLIVAN, J. M., PANADES, M., PARKER, C., PATTERSON, H., SCRASE, C., STAFFURTH, J., STOCKDALE, A., TREMLETT, J., BIDMEAD, M., MAYLES, H., NAISMITH, O., SOUTH, C., GAO, A., CRUICKSHANK, C., HASSAN, S., PUGH, J., GRIFFIN, C. & HALL, E. 2016. Conventional versus hypofractionated high-dose intensity-modulated radiotherapy for prostate cancer: 5-year outcomes of the randomised, non-inferiority, phase 3 CHHiP trial. *The Lancet Oncology,* 17**,** 1047-1060.

DEARNALEY, D. P., KHOO, V. S., NORMAN, A. R., MEYER, L., NAHUM, A., TAIT, D., YARNOLD, J. & HORWICH, A. 1999. Comparison of radiation side-effects of conformal and conventional radiotherapy in prostate cancer: a randomised trial. *Lancet,* 353**,** 267-272.

DENMEADE, S. R. & ISAACS, J. T. 2004. Development of prostate cancer treatment: The good news. *Prostate,* 58**,** 211-224.

DIEDERICH, J., RUHMANN, I. & MAY, M. 1987. KRITON: a knowledge-acquisition tool for expert systems. *International Journal of Man-Machine Studies,* 26**,** 29-40.

DJAVAN, B., FONG, Y. K., HARIK, M., MILANI, S., REISSIGL, A., CHAUDRY, A., ANAGNOSTOU, T., BAGHERI, F., WALDERT, M., KREUZER, S., FAJKOVIC, H. & MARBERGER, M. 2004. Longitudinal study of men with mild symptoms of bladder outlet obstruction treated with watchful waiting for four years. *Urology,* 64**,** 1144-1148.

DJAVAN, B., REMZI, M., ZLOTTA, A., SEITZ, C., SNOW, P. & MARBERGER, M. 2002. Novel artificial neural network for early detection of prostate cancer. *Journal of Clinical Oncology,* 20**,** 921-929.

DOUST, J., MILLER, E., DUCHESNE, G., KITCHENER, M. & WELLER, D. 2004. A systematic review of brachytherapy. Is it an effective and safe treatment for localised prostate cancer? *Australian family physician,* 33**,** 525-9.

EAU, N. M. C., J. BELLMUNT, E. BRIERS (PATIENT, REPRESENTATIVE), R. C. N. V. D. B. G. A., M. BOLLA, N. J. V. C. G. A., P. CORNFORD,, S. CULINE, S. J., T. LAM, M.D. MASON, V. MATVEEV, & H. VAN DER POEL, T. H. V. D. K., O. ROUVIÈRE, T. WIEGEL. 2015. *Guidelines on Prostate Cancer* [Online]. Available: https://uroweb.org/wp-content/uploads/EAU-Guidelines-Prostate-Cancer-2015-v2.pdf.

ECKE, T. H., BARTEL, P., HALLMANN, S., KOCH, S., RUTTLOFF, J., CAMMANN, H., LEIN, M., SCHRADER, M., MILLER, K. & STEPHAN, C. 2012. Outcome prediction for prostate cancer detection rate with artificial neural network (ANN) in daily routine. *Urologic Oncology-Seminars and Original Investigations,* 30**,** 139-144.

EL-MEKRESH, M., AKL, A., MOSBAH, A., ABDEL-LATIF, M., ABOL-ENEIN, H. & GHONEIM, M. A. 2009. Prediction of Survival After Radical Cystectomy for Invasive Bladder Carcinoma: Risk Group Stratification, Nomograms or Artificial Neural Networks? *The Journal of Urology,* 182**,** 466-472.

ENTING, J., HUIRNE, R. B. M., DIJKHUIZEN, A. A. & TIELEN, M. J. M. 1999. A knowledge documentation methodology for knowledge-based system development: an example in animal health management. *Computers and Electronics in Agriculture,* 22**,** 117-129.

FAITHFULL, S., CORNER, J., MEYER, L., HUDDART, R. & DEARNALEY, D. 2001. Evaluation of nurse-led follow up for patients undergoing pelvic radiotherapy. *British Journal of Cancer,* 85**,** 1853-1864.

FILELLA, X., FOJ, L., ALCOVER, J., AUGÉ, J. M., MOLINA, R. & JIMÉNEZ, W. 2014. The influence of prostate volume in prostate health index performance in patients with total PSA lower than 10 μg/L. *Clinica Chimica Acta,* 436**,** 303-307.

FINNE, P., FINNE, R., AUVINEN, A., JUUSELA, H., ARO, J., MAATTANEN, L., HAKAMA, M., RANNIKKO, S., TAMMELA, T. L. J. & STENMAN, U. H. 2000. Predicting the outcome of prostate biopsy in screen-positive men by a multilayer perceptron network. *Urology,* 56**,** 418-422.

FOSSATI, N., KARNES, R. J., COZZARINI, C., FIORINO, C., GANDAGLIA, G., JONIAU, S., BOORJIAN, S. A., GOLDNER, G., HINKELBEIN, W., HAUSTERMANS, K., TOMBAL, B., SHARIAT, S., KARAKIEWICZ, P. I., MONTORSI, F., VAN POPPEL, H., WIEGEL, T. & BRIGANTI, A. 2016. Assessing the Optimal Timing for Early Salvage Radiation Therapy in Patients with Prostate-specific Antigen Rise After Radical Prostatectomy. *European Urology,* 69**,** 728-733.

FREEDLAND, S. J., HUMPHREYS, E. B., MANGOLD, L. A., EISENBERGER, M., DOREY, F. J., WALSH, P. C. & PARTIN, A. W. 2005. Risk of prostate cancer-specific mortality following biochemical recurrence after radical prostatectomy. *Jama-Journal of the American Medical Association,* 294**,** 433-439.

FUJIHARA, H., SIMMONS, D. B., ELLIS, N. C. & SHANNON, R. E. 1997. Knowledge conceptualization tool. *Ieee Transactions on Knowledge and Data Engineering,* 9**,** 209-220.

FUJIKAWA, K., MATSUI, Y., KOBAYASHI, T., MIURA, K., OKA, H., FUKUZAWA, S., SASAKI, M., TAKEUCHI, H. & OKABE, T. 2003. Predicting disease outcome of non-invasive transitional cell carcinoma of the urinary bladder using an artificial neural network model: Results of patient follow-up for 15 years or longer. *International Journal of Urology,* 10**,** 149-152.

GAINES, B. R. & SHAW, M. L. G. 1993. ELICITING KNOWLEDGE AND TRANSFERRING IT EFFECTIVELY TO A KNOWLEDGE-BASED SYSTEM. *Ieee Transactions on Knowledge and Data Engineering,* 5**,** 4-14.

GARG, A. X., ADHIKARI, N. K. J., MCDONALD, H., ROSAS-ARELLANO, M. P., DEVEREAUX, P. J., BEYENE, J., SAM, J. & HAYNES, R. B. 2005. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes - A systematic review. *Jama-Journal of the American Medical Association,* 293**,** 1223-1238.

GATIDIS, S., SCHARPF, M., MARTIROSIAN, P., BEZRUKOV, I., KÜSTNER, T., HENNENLOTTER, J., KRUCK, S., KAUFMANN, S., SCHRAML, C. & FOUGÈRE, C. 2015. Combined unsupervised–supervised classification of multiparametric PET/MRI data: application to prostate cancer. *NMR in Biomedicine,* 28**,** 914-922.

GEBUS, S. & LEIVISKA, K. 2009. Knowledge acquisition for decision support systems on an electronic assembly line. *Expert Systems with Applications,* 36**,** 93-101.

GIL, D., GIRELA, J. L., DE JUAN, J., GOMEZ-TORRES, M. J. & JOHNSSON, M. 2012. Predicting seminal quality with artificial intelligence methods. *Expert Systems with Applications,* 39**,** 12564-12573.

GIL, D. & JOHNSSON, M. 2010. Using support vector machines in diagnoses of urological dysfunctions. *Expert Systems with Applications,* 37**,** 4713-4718.

GIL, D., JOHNSSON, M., CHAMIZO, J. M. G., PAYA, A. S. & FERNANDEZ, D. R. 2009. Application of artificial neural networks in the diagnosis of urological dysfunctions. *Expert Systems with Applications,* 36**,** 5754-5760.

GIRELA, J. L., GIL, D., JOHNSSON, M., GOMEZ-TORRES, M. J. & DE JUAN, J. 2013. Semen Parameters Can Be Predicted from Environmental Factors and Lifestyle Using Artificial Intelligence Methods. *Biology of Reproduction,* 88.

GOMHA, M. A., SHEIR, K. Z., SHOWKY, S., ABDEL-KHALEK, M., MOKHTAR, A. A. & MADBOULY, K. 2004. Can we improve the prediction of stone-free status after extracorporeal shock wave lithotripsy for ureteral stones? A neural network or a statistical model? *J Urol,* 172**,** 175-9.

GORMAN, R. 1995. Expert system for management of urinary incontinence in women. *Proc Annu Symp Comput Appl Med Care***,** 527-31.

GOYAL, N. K., KUMAR, A., ACHARYA, R. L., DWIVEDI, U. S., TRIVEDI, S., SINGH, P. B. & SINGH, T. N. 2007. Prediction of biochemical failure in localized carcinoma of prostate after radical prostatectomy by neuro-fuzzy. *Indian J Urol,* 23**,** 14-7.

GREEN, W. J. F., BALL, G., HULMAN, G., JOHNSON, C., VAN SCHALWYK, G., RATAN, H. L., SORIA, D., GARIBALDI, J. M., PARKINSON, R., HULMAN, J., REES, R. & POWE, D. G. 2016. KI67 and DLX2 predict increased risk of metastasis formation in prostate cancer-a targeted molecular approach. *Br J Cancer,* 115**,** 236-242.

GROSAN, C. & ABRAHAM, A. 2011. Rule-Based Expert Systems. *Intelligent Systems: a Modern Approach.* Berlin: Springer-Verlag Berlin.

HAMID, A., DWIVEDI, U. S., SINGH, T. N., KISHORE, M. G., MAHMOOD, M., SINGH, H., TANDON, V. & SINGH, P. B. 2003. Artificial neural networks in predicting optimum renal stone fragmentation by extracorporeal shock wave lithotripsy: a preliminary study. *Bju International,* 91**,** 821-824.

HAN, M., PARTIN, A. W., POUND, C. R., EPSTEIN, J. I. & WALSH, P. C. 2001a. Long-term biochemical disease-free and cancer-specific survival following anatomic radical retropubic prostatectomy - The 15-year Johns Hopkins experience. *Urologic Clinics of North America,* 28**,** 555-+.

HAN, M., SNOW, P. B., BRANDT, J. M. & PARTIN, A. W. 2001b. Evaluation of artificial neural networks for the prediction of pathologic stage in prostate carcinoma. *Cancer,* 91**,** 1661-1666.

HAN, M., SNOW, P. B., EPSTEIN, J. I., CHAN, T. Y., JONES, K. A., WALSH, P. C. & PARTIN, A. W. 2000. A neural network predicts progression for men with Gleason score 3+4 versus 4+3 tumors after radical prostatectomy. *Urology,* 56**,** 994-999.

HANCHANALE, V. S., MCCABE, J. E. & JAVLE, P. 2010. Radical Prostatectomy Practice in England. *Urology Journal,* 7**,** 243-248.

HAO, A. T. H., WU, L. P., KUMAR, A., JIAN, W. S., HUANG, L. F., KAO, C. C. & HSU, C. Y. 2013. Nursing process decision support system for urology ward. *International Journal of Medical Informatics,* 82**,** 604-612.

HASSANIEN, A. E., AL-QAHERI, H. & EL-DAHSHAN, E. S. A. 2011. Prostate boundary detection in ultrasound images using biologically-inspired spiking neural network. *Applied Soft Computing,* 11**,** 2035-2041.

HEEMSBERGEN, W. D., AL-MAMGANI, A., SLOT, A., DIELWART, M. F. H. & LEBESQUE, J. V. 2014. Long-term results of the Dutch randomized prostate cancer trial: Impact of dose-escalation on local, biochemical, clinical failure, and survival. *Radiotherapy and Oncology,* 110**,** 104-109.

HELGESEN, F., ANDERSSON, S. O., GUSTAFSSON, O., VARENHORST, E., GOBEN, B., CARNOCK, S., SEHLSTEDT, L., CARLSSON, P., HOLMBERG, L. & JOHANSSON, J. E. 2000. Follow-up of prostate cancer patients by on-demand contacts with a specialist nurse - A randomized study. *Scandinavian Journal of Urology and Nephrology,* 34**,** 55-61.

HEMMINKI, K. & CHEN, B. W. 2005. Familial association of prostate cancer with other cancers in the Swedish family-cancer database. *Prostate,* 65**,** 188-194.

HOLM, H. H., JUUL, N., PEDERSEN, J. F., HANSEN, H. & STROYER, I. 1983. TRANSPERINEAL I125 SEED IMPLANTATION IN PROSTATIC-CANCER GUIDED BY TRANS-RECTAL ULTRASONOGRAPHY. *Journal of Urology,* 130**,** 283-286.

HOLZBEIERLEIN, J. M., LANGENSTROER, P., PORTER, H. J., 2ND & THRASHER, J. B. 2003. Case selection and outcome of radical perineal prostatectomy in localized prostate cancer. *International braz j urol : official journal of the Brazilian Society of Urology,* 29**,** 291-9.

HORWITZ, E. M., THAMES, H. D., KUBAN, D. A., LEVY, L. B., KUPELIAN, P. A., MARTINEZ, A. A., MICHALSKI, J. M., PISANSKY, T. M., SANDLER, H. M., SHIPLEY, W. U., ZELEFSKY, M. J., HANKS, G. E. & ZIETMAN, A. L. 2005. 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. *Journal of Urology,* 173**,** 797-802.

HOSKIN, P. J., COLOMBO, A., HENRY, A., NIEHOFF, P., PAULSEN HELLEBUST, T., SIEBERT, F.-A. & KOVACS, G. 2013. GEC/ESTRO recommendations on high dose rate afterloading brachytherapy for localised prostate cancer: An update. *Radiotherapy and Oncology,* 107**,** 325-332.

HOSKIN, P. J., ROJAS, A. M., BOWNES, P. J., LOWE, G. J., OSTLER, P. J. & BRYANT, L. 2012. Randomised trial of external beam radiotherapy alone or combined with high-dose-rate brachytherapy boost for localised prostate cancer. *Radiother Oncol,* 103**,** 217-22.

HU, X. H., CAMMANN, H., MEYER, H. A., JUNG, K., LU, H. B., LEVA, N., MAGHELI, A., STEPHAN, C. & BUSCH, J. 2014. Risk prediction models for biochemical recurrence after radical prostatectomy using prostate-specific antigen and Gleason score. *Asian Journal of Andrology,* 16**,** 897-901.

HUGGINS, C. & HODGES, C. V. 1941. Studies on prostatic cancer - I The effect of castration, of estrogen and of androgen injection on serum phosphatases in metastatic carcinoma of the prostate. *Cancer Research,* 1**,** 293-297.

HUGGINS, C., STEVENS, R. E. & HODGES, C. V. 1941. Studies on prostate cancer II The effects of castration on advanced carcinoma of the prostate gland. *Archives of Surgery,* 43**,** 209-223.

HUMMEL, S., PAISLEY, S., MORGAN, A., CURRIE, E. & BREWER, N. 2003. Clinical and cost-effectiveness of new and emerging technologies for early localised prostate cancer: a systematic review. *Health technology assessment (Winchester, England),* 7**,** ix-x, 1-157.

HUNTER, J., ; OTTLEY, DREWRY; PALMER, JAMES F. ED; BELL, THOMAS, ; BABINGTON, GEORGE GISBORNE, HOME, EVERARD, SIR, OWEN, RICHARD, MAILE, G., ENGRAVER; SHARP, W., ENGRAVER; SKELTON, W., ENGRAVER; ROBERTS, J., ENGRAVER; PINX, REYNOLDS, ARTIST; BELL, WILLIAM, ARTIST; AUBIN, S., ARTIST; RIEMSDYKE, J. V., ARTIST; STAURT, J., ARTIST; BASIRE, J., LITHOGRAPHER; SCHARF, G., LITHOGRAPHER; TAYLOR, RICHARD, PRINTER 1837. *The works of John Hunter F.R.S with notes*, LONGMAN.

HURST, R. E., BONNER, R. B., ASHENAYI, K., VELTRI, R. W. & HEMSTREET, G. P., 3RD 1997. Neural net-based identification of cells expressing the p300 tumor-related antigen using fluorescence image analysis. *Cytometry,* 27**,** 36-42.

ILIC, D., EVANS, S. M., ALLAN, C. A., JUNG, J. H., MURPHY, D. & FRYDENBERG, M. 2017. Laparoscopic and robotic-assisted versus open radical prostatectomy for the treatment of localised prostate cancer. *Cochrane Database Syst Rev,* 9**,** CD009625.

JACKSON, P. 1999. *Introduction to Expert Systems*, Addison-Wesley.

JAMES, N. D., DE BONO, J. S., SPEARS, M. R., CLARKE, N. W., MASON, M. D., DEARNALEY, D. P., RITCHIE, A. W. S., AMOS, C. L., GILSON, C., JONES, R. J., MATHESON, D., MILLMAN, R., ATTARD, G., CHOWDHURY, S., CROSS, W. R., GILLESSEN, S., PARKER, C. C., RUSSELL, J. M., BERTHOLD, D. R., BRAWLEY, C., ADAB, F., AUNG, S., BIRTLE, A. J., BOWEN, J., BROCK, S., CHAKRABORTI, P., FERGUSON, C., GALE, J., GRAY, E., HINGORANI, M., HOSKIN, P. J., LESTER, J. F., MALIK, Z. I., MCKINNA, F., MCPHAIL, N., MONEY-KYRLE, J., O'SULLIVAN, J., PARIKH, O., PROTHEROE, A., ROBINSON, A., SRIHARI, N. N., THOMAS, C., WAGSTAFF, J., WYLIE, J., ZARKAR, A., PARMAR, M. K. B. & SYDES, M. R. 2017. Abiraterone for Prostate Cancer Not Previously Treated with Hormone Therapy. *N Engl J Med,* 377**,** 338-351.

JAMES, N. D., SYDES, M. R., CLARKE, N. W., MASON, M. D., DEARNALEY, D. P., SPEARS, M. R., RITCHIE, A. W. S., PARKER, C. C., RUSSELL, J. M., ATTARD, G., DE BONO, J., CROSS, W., JONES, R. J., THALMANN, G., AMOS, C., MATHESON, D., MILLMAN, R., ALZOUEBI, M., BEESLEY, S., BIRTLE, A. J., BROCK, S., CATHOMAS, R., CHAKRABORTI, P., CHOWDHURY, S., COOK, A., ELLIOTT, T., GALE, J., GIBBS, S., GRAHAM, J. D., HETHERINGTON, J., HUGHES, R., LAING, R., MCKINNA, F., MCLAREN, D. B., O'SULLIVAN, J. M., PARIKH, O., PEEDELL, C., PROTHEROE, A., ROBINSON, A. J., SRIHARI, N., SRINIVASAN, R., STAFFURTH, J., SUNDAR, S., TOLAN, S., TSANG, D., WAGSTAFF, J. & PARMAR, M. K. B. 2016. Addition of docetaxel, zoledronic acid, or both to first-line long-term hormone therapy in prostate cancer (STAMPEDE): survival results from an adaptive, multiarm, multistage, platform randomised controlled trial. *The Lancet,* 387**,** 1163-1177.

JOHNS, L. E. & HOULSTON, R. S. 2003. A systematic review and meta-analysis of familial prostate cancer risk. *Bju International,* 91**,** 789-794.

JOHNSON, R. B. & ONWUEGBUZIE, A. J. 2004. Mixed Methods Research: A Research Paradigm Whose Time Has Come. *Educational Researcher,* 33**,** 14-26.

JONIAU, S., BRIGANTI, A., GONTERO, P., GANDAGLIA, G., TOSCO, L., FIEUWS, S., TOMBAL, B., MARCHIORO, G., WALZ, J., KNEITZ, B., BADER, P., FROHNEBERG, D., TIZZANI, A., GRAEFEN, M., VAN CANGH, P., KARNES, R. J., MONTORSI, F., VAN POPPEL, H. & SPAHN, M. 2015. Stratification of High-risk Prostate Cancer into Prognostic Categories: A European Multi-institutional Study. *European Urology,* 67**,** 157-164.

KALRA, P., TOGAMI, J., BANSAL, G., PARTIN, A. W., BRAWER, M. K., BABAIAN, R. J., ROSS, L. S. & NIEDERBERGER, C. S. 2003. A neurocomputational model for prostate carcinoma detection. *Cancer,* 98**,** 1849-1854.

KASIVISVANATHAN, V., RANNIKKO, A. S., BORGHI, M., PANEBIANCO, V., MYNDERSE, L. A., VAARALA, M. H., BRIGANTI, A., BUDÄUS, L., HELLAWELL, G., HINDLEY, R. G., ROOBOL, M. J., EGGENER, S., GHEI, M., VILLERS, A., BLADOU, F., VILLEIRS, G. M., VIRDI, J., BOXLER, S., ROBERT, G., SINGH, P. B., VENDERINK, W., HADASCHIK, B. A., RUFFION, A., HU, J. C., MARGOLIS, D., CROUZET, S., KLOTZ, L., TANEJA, S. S., PINTO, P., GILL, I., ALLEN, C., GIGANTI, F., FREEMAN, A., MORRIS, S., PUNWANI, S., WILLIAMS, N. R., BREW-GRAVES, C., DEEKS, J., TAKWOINGI, Y., EMBERTON, M., MOORE, C. M., KENTTÄMIES, A., MIRTTI, T., BECHER, E. F., CATALANO, C., GROMPONE, M., DEL MONTE, M., COSTANTINO, L., SCIARRA, A., D'ERAMO, G., SALVO, V., CAMPA, R., KING, B. F., FROEMMING, A. T., MCLAREN, R. H., DRAAYER, P. J., SMITH, J. E., DOTY, K. J., TONTTILA, P., VIRTA, V., KUISMA, M., PÄÄKKÖ, E., HIRVIKOSKI, P., MONTORSI, F., STABILE, A., DE COBELLI, F., ESPOSITO, A., PICOZZI, M., CRISTEL, G., BREMBILLA, G., HUNG, M., OCAMPO, F., OTIENO, M., RAMACHANDRAN, N., PENDSE, D., KIRKHAM, A., JAMESON, C., RATYNSKA, M., BEN-SALHA, I., LEYH-BANNURAH, S. R., BEYERSDORFF, D., SAUTER, G., CHAMBERS, A., PATHMANATHAN, B., GAYED, W., VRENTZOU, E., BALDWIN, R., RAJKUMAR, G., EMARA, A., NEDAS, T., EDWARDS, A., BANGMA, C. H., BUSSTRA, M. B., SCHOOTS, I. G., OTO, A., GERBER, G., SHAH, T., HAMID, S., EROTOCRITOU, P., MARAJ, B., et al. 2018. MRI-targeted or standard biopsy for prostate-cancer diagnosis. *New England Journal of Medicine,* 378**,** 1767-1777.

KATTAN, M. W., COWEN, M. E. & MILES, B. J. 1996. Computer modeling in urology. *Urology,* 47**,** 14-21.

KAWAKAMI, S., NUMAO, N., OKUBO, Y., KOGA, F., YAMAMOTO, S., SAITO, K., FUJII, Y., YONESE, J., MASUDA, H., KIHARA, K. & FUKUI, I. 2008. Development, validation, and head-to-head comparison of logistic regression-based nomograms and artificial neural network models predicting prostate cancer on initial extended biopsy. *European Urology,* 54**,** 601-611.

KAWAMOTO, K., HOULIHAN, C. A., BALAS, E. A. & LOBACH, D. F. 2005. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. *BMJ,* 330**,** 765.

KELES, A., HASILOGLU, A. S., KELES, A. & AKSOY, Y. 2007. Neuro-fuzzy classification of prostate cancer using NEFCLASS-J. *Computers in Biology and Medicine,* 37**,** 1617-1628.

KICINSKI, M., VANGRONSVELD, J. & NAWROT, T. S. 2011. An Epidemiological Reappraisal of the Familial Aggregation of Prostate Cancer: A Meta-Analysis. *Plos One,* 6**,** 7.

KIDD, A. 1987. Knowledge Acquisition. *In:* KIDD, A. (ed.) *Knowledge Acquisition for Expert Systems.* Springer US.

KIM, M., RAMIREZ, L., YOO, C., CHOO, M., PAICK, J.-S. & OH, S.-J. 2014. Factors Influencing Nonabsolute Indications for Surgery in Patients With Lower Urinary Tract Symptoms Suggestive of Benign Prostatic Hyperplasia: Analysis Using Causal Bayesian Networks. *Int Neurourol J,* 18**,** 198-205.

KIM, S. Y., MOON, S. K., JUNG, D. C., HWANG, S. I., SUNG, C. K., CHO, J. Y., KIM, S. H., LEE, J. & LEE, H. J. 2011. Pre-Operative Prediction of Advanced Prostatic Cancer Using Clinical Decision Support Systems: Accuracy Comparison between Support Vector Machine and Artificial Neural Network. *Korean Journal of Radiology,* 12**,** 588-594.

KNAUF, R., GONZALEZ, A. J. & ABEL, T. 2002. A framework for validation of rule-based systems. *Ieee Transactions on Systems Man and Cybernetics Part B-Cybernetics,* 32**,** 281-295.

KOLASA, M., WOJTYNA, R., DLUGOSZ, R. & JOZWICKI, W. 2009. Application of Artificial Neural Network to Predict Survival Time for Patients with Bladder Cancer. *Computers in Medical Activity,* 65**,** 113-122.

KOUTSOJANNIS, C. & HATZILYGEROUDIS, L. 2004. FESMI: A fuzzy expert system for diagnosis and treatment of male impotence. *Knowledge-Based Intelligent Information and Engineering Systems, Pt 2, Proceedings,* 3214**,** 1106-1113.

KOUTSOJANNIS, C., LITHARI, C. & HATZILYGEROUDIS, I. 2012. Managing urinary incontinence through hand-held real-time decision support aid. *Computer Methods and Programs in Biomedicine,* 107**,** 84-89.

KOUTSOJANNIS, C., NABIL, E., TSIMARA, M., HATZILYGEROUDIS, I. & IEEE 2009. Using Machine Learning Techniques to improve the behaviour of a medical decision support system for prostate diseases. *2009 9th International Conference on Intelligent Systems Design and Applications***,** 341-346.

KOUTSOJANNIS, C., TSIMARA, M. & NABIL, E. 2008. *HIROFILOS: A Medical Expert System for Prostate Diseases*.

KRONGRAD, A., GRANVILLE, L. J., BURKE, M. A., GOLDEN, R. M., LAI, S. H., CHO, L. & NIEDERBERGER, C. S. 1997. Predictors of general quality of life in patients with benign prostate hyperplasia or prostate cancer. *Journal of Urology,* 157**,** 534-538.

KSHIRSAGAR, A., SEFTEL, A., ROSS, L., MOHAMED, M. & NIEDERBERGER, C. 2006. Predicting hypogonadism in men based upon age, presence of erectile dysfunction, and depression. *International Journal of Impotence Research,* 18**,** 47-51.

KUBAN, D. A., LEVY, L. B., POTTERS, L., BEYER, D. C., BLASKO, J. C., MORAN, B. J., CIEZKI, J. P., ZIETMAN, A. L., ZELEFSKY, M. J., PISANSKY, T. M., ELSHAIKH, M. & HORWITZ, E. M. 2006. Comparison of biochemical failure definitions for permanent prostate brachytherapy. *International Journal of Radiation Oncology Biology Physics,* 65**,** 1487-1493.

KUMAR, S., SHELLEY, M., HARRISON, C., COLES, B., WILT, T. J. & MASON, M. 2006. Neo-adjuvant and adjuvant hormone therapy for localised and locally advanced prostate cancer. *Cochrane Database of Systematic Reviews* [Online]. Available: http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD006019.pub2/abstract.

KUO, R.-J., HUANG, M.-H., CHENG, W.-C., LIN, C.-C. & WU, Y.-H. 2015. Application of a two-stage fuzzy neural network to a prostate cancer prognosis system. *Artificial Intelligence in Medicine,* 63**,** 119-133.

LAMB, D. J. & NIEDERBERGER, C. S. 1993. ARTIFICIAL-INTELLIGENCE IN MEDICINE AND MALE-INFERTILITY. *World Journal of Urology,* 11**,** 129-136.

LAWRENTSCHUK, N., LOCKWOOD, G., DAVIES, P., EVANS, A., SWEET, J., TOI, A. & FLESHNER, N. E. 2011. Predicting prostate biopsy outcome: artificial neural networks and polychotomous regression are equivalent models. *International Urology and Nephrology,* 43**,** 23-30.

LEE, H. J., HWANG, S. I., HAN, S. M., PARK, S. H., KIM, S. H., CHO, J. Y., SEONG, C. G. & CHOE, G. 2010. Image-based clinical decision support for transrectal ultrasound in the diagnosis of prostate cancer: comparison of multiple logistic regression, artificial neural network, and support vector machine. *European Radiology,* 20**,** 1476-1484.

LEE, H. J., KIM, K. G., LEE, S. E., BYUN, S. S., HWANG, S. I., JUNG, S. I., HONG, S. K. & KIM, S. H. 2006. Role of transrectal ultrasonography in the prediction of prostate cancer - Artificial neural network analysis. *Journal of Ultrasound in Medicine,* 25**,** 815-821.

LEE, S. & OKEEFE, R. M. 1994. DEVELOPING A STRATEGY FOR EXPERT-SYSTEM VERIFICATION AND VALIDATION. *Ieee Transactions on Systems Man and Cybernetics,* 24**,** 643-655.

LIAO, S. H. 2005. Expert system methodologies and applications - a decade review from 1995 to 2004. *Expert Systems with Applications,* 28**,** 93-103.

LIU, H., HEMMINKI, K. & SUNDQUIST, J. 2011. Renal Cell Carcinoma as First and Second Primary Cancer: Etiological Clues From the Swedish Family-Cancer Database. *Journal of Urology,* 185**,** 2045-2049.

LLOBET, R., PEREZ-CORTES, J. C., TOSELLI, A. H. & JUAN, A. 2007. Computer-aided detection of prostate cancer. *International Journal of Medical Informatics,* 76**,** 547-556.

LOCH, T., LEUSCHNER, I., GENBERG, C., WEICHERT-JACOBSEN, K., KUPPERS, F., YFANTIS, E., EVANS, M., TSAREV, V. & STOCKLE, M. 1999. Artificial neural network analysis (ANNA) of prostatic transrectal ultrasound. *Prostate,* 39**,** 198-204.

LOGVINENKO, T., CHOW, J. S. & NELSON, C. P. 2015. Predictive value of specific ultrasound findings when used as a screening test for abnormalities on VCUG. *Journal of Pediatric Urology,* 11**,** 176.e1-176.e7.

LOPES, M., ORTEGA, N. R. S., SILVEIRA, P. S. P., MASSAD, E., HIGA, R. & MARIN, H. D. 2013. Fuzzy cognitive map in differential diagnosis of alterations in urinary elimination: A nursing approach. *International Journal of Medical Informatics,* 82**,** 201-208.

LUENGO-FERNANDEZ, R., LEAL, J., GRAY, A. & SULLIVAN, R. 2013. Economic burden of cancer across the European Union: a population-based cost analysis. *The Lancet Oncology,* 14**,** 1165-1174.

LUGER, G. F. 1998. Knowledge intensive problem solving. *In:* STUBBLFIELD, W. A. (ed.) *Artificial Intelligence: Structures and Strategies for Complex Problem Solving.* Addison-Wesley, Reading, MA.

LUNDSGAARDE, H. P. 1987. Evaluating medical expert systems. *Social Science & Medicine,* 24**,** 805-819.

MACDONALD, B. A. & WITTEN, I. H. 1989. A FRAMEWORK FOR KNOWLEDGE ACQUISITION THROUGH TECHNIQUES OF CONCEPT-LEARNING. *Ieee Transactions on Systems Man and Cybernetics,* 19**,** 499-512.

MARSZALL, M. P., KRYSINSKI, J., SROKA, W. D., NYCZAK, Z., STEFANOWICZ, M., WASNIEWSKI, T., ROMASZKO, J. & BUCINSKI, A. 2012. ANN as a prognostic tool after treatment of non-seminoma testicular cancer. *Central European Journal of Medicine,* 7**,** 672-679.

MASON, M. D., PARULEKAR, W. R., SYDES, M. R., BRUNDAGE, M., KIRKBRIDE, P., GOSPODAROWICZ, M., COWAN, R., KOSTASHUK, E. C., ANDERSON, J., SWANSON, G., PARMAR, M. K. B., HAYTER, C., JOVIC, G., HILTZ, A., HETHERINGTON, J., SATHYA, J., BARBER, J. B. P., MCKENZIE, M., EL-SHARKAWI, S., SOUHAMI, L., HARDMAN, P. D. J., CHEN, B. E. & WARDE, P. 2015. Final Report of the Intergroup Randomized Study of Combined Androgen-Deprivation Therapy Plus Radiotherapy Versus Androgen-Deprivation Therapy Alone in Locally Advanced Prostate Cancer. *Journal of Clinical Oncology,* 33**,** 2143-2150.

MATSUI, Y., EGAWA, S., TSUKAYAMA, C., TERAI, A., KUWAO, S., BABA, S. & ARAI, Y. 2002. Artificial neural network analysis for predicting pathological stage of clinically localized prostate cancer in the Japanese population. *Japanese Journal of Clinical Oncology,* 32**,** 530-535.

MATSUI, Y., UTSUNOMIYA, N., ICHIOKA, K., UEDA, N., YOSHIMURA, K., TERAI, A. & ARAI, Y. 2004. The use of artificial neural network analysis to improve the predictive accuracy of prostate biopsy in the Japanese population. *Japanese Journal of Clinical Oncology,* 34**,** 602-607.

MATTFELDT, T., KESTLER, H. A., HAUTMANN, R. & GOTTFRIED, H. W. 1999. Prediction of prostatic cancer progression after radical prostatectomy using artificial neural networks: a feasibility study. *Bju International,* 84**,** 316-323.

MATTFELDT, T., KESTLER, H. A., HAUTMANN, R. & GOTTFRIED, H. W. 2001. Prediction of postoperative prostatic cancer stage on the basis of systematic biopsies using two types of artificial neural networks. *European Urology,* 39**,** 530-536.

MATULEWICZ, L., JANSEN, J. F. A., BOKACHEVA, L., VARGAS, H. A., AKIN, O., FINE, S. W., SHUKLA-DAVE, A., EASTHAM, J. A., HRICAK, H., KOUTCHER, J. A. & ZAKIAN, K. L. 2014. Anatomic Segmentation Improves Prostate Cancer Detection With Artificial Neural Networks Analysis of H-1 Magnetic Resonance Spectroscopic Imaging. *Journal of Magnetic Resonance Imaging,* 40**,** 1414-1421.

MCCARTHY, J., MINSKY, M. L. & SHANNON, C. E. 2006. A proposal for the Dartmouth summer research project on artificial intelligence - August 31, 1955. *Ai Magazine,* 27**,** 12-14.

MCINTOSH, H. M., NEAL, R. D., ROSE, P., WATSON, E., WILKINSON, C., WELLER, D., CAMPBELL, C. & PROSTATE CANC FOLLOW-UP, G. 2009. Follow-up care for men with prostate cancer and the role of primary care: a systematic review of international guidelines. *British Journal of Cancer,* 100**,** 1852-1860.

MCVEY, G. P., MCPHAIL, S., FOWLER, S., MCINTOSH, G., GILLATT, D. & PARKER, C. C. 2010. Initial management of low-risk localized prostate cancer in the UK: analysis of the British Association of Urological Surgeons Cancer Registry. *Bju International,* 106**,** 1161-1164.

MEIJER, R. P., GEMEN, E. F. A., VAN ONNA, I. E. W., VAN DER LINDEN, J. C., BEERLAGE, H. P. & KUSTERS, G. C. M. 2009. The value of an artificial neural network in the decision-making for prostate biopsies. *World Journal of Urology,* 27**,** 593-598.

MICHAELS, E. K., NIEDERBERGER, C. S., GOLDEN, R. M., BROWN, B., CHO, L. & HONG, Y. 1998. Use of a neural network to predict stone growth after shock wave lithotripsy. *Urology,* 51**,** 335-338.

MILLIN, T. 1945. Retropubic prostatectomy; a new extravesical technique; report of 20 cases. *Lancet,* 2**,** 693-6.

MISTRY, M., PARKIN, D. M., AHMAD, A. S. & SASIENI, P. 2011. Cancer incidence in the United Kingdom: projections to the year 2030. *British Journal of Cancer,* 105**,** 1795-1803.

MITRA, A., FISHER, C., FOSTER, C. S., JAMESON, C., BARBACHANNO, Y., BARTLETT, J., BANCROFT, E., DOHERTY, R., KOTE-JARAI, Z., PEOCK, S., EASTON, D. & EELES, R. 2008. Prostate cancer in male BRCA1 and BRCA2 mutation carriers has a more aggressive phenotype. *Br J Cancer,* 98**,** 502-7.

MORRIS, W. J., TYLDESLEY, S., RODDA, S., HALPERIN, R., PAI, H., MCKENZIE, M., DUNCAN, G., MORTON, G., HAMM, J. & MURRAY, N. 2017. Androgen Suppression Combined with Elective Nodal and Dose Escalated Radiation Therapy (the ASCENDE-RT Trial): An Analysis of Survival Endpoints for a Randomized Trial Comparing a Low-Dose-Rate Brachytherapy Boost to a Dose-Escalated External Beam Boost for High- and Intermediate-risk Prostate Cancer. *International Journal of Radiation Oncology*Biology*Physics,* 98**,** 275-285.

MOTTET, N., BELLMUNT, J., BRIERS, E., BOLLA, M., BOURKE, L., CORNFORD, P., DE SANTIS, M., HENRY, A., JONIAU, S. & LAM, T. 2017. EAU-ESTRO-ESUR-SIOG Guidelines on.

MOUL, J. W., SNOW, P. B., FERNANDEZ, E. B., MAHER, P. D. & SESTERHENN, I. A. 1995. NEURAL-NETWORK ANALYSIS OF QUANTITATIVE HISTOLOGICAL FACTORS TO PREDICT PATHOLOGICAL STAGE IN CLINICAL STAGE-I NONSEMINOMATOUS TESTICULAR CANCER. *Journal of Urology,* 153**,** 1674-1677.

NAGUIB, R. N. G., QURESHI, K. N., HAMDY, F. C. & NEAL, D. E. Neural network analysis of prognostic markers in bladder cancer. Engineering in Medicine and Biology Society, 1997. Proceedings of the 19th Annual International Conference of the IEEE, 30 Oct-2 Nov 1997 1997. 1007-1009 vol.3.

NAGUIB, R. N. G., ROBINSON, M. C., NEAL, D. E. & HAMDY, F. C. 1998. Neural network analysis of combined conventional and experimental prognostic markers in prostate cancer: a pilot study. *British Journal of Cancer,* 78**,** 246-250.

NCIN. 2009. *Cancer Incidence and Survival by Major Ethnic Group, England 2002-2006* [Online]. National Cancer Intelligence Network and Cancer Research UK. Available: http://www.ncin.org.uk/view.aspx?rid=75 2009].

NCPA 2014. NCPA 2014 Annual Report. http://www.npca.org.uk/reports/: National Prostate Cancer Audit.

NEUZILLET, Y., LECHEVALLIER, E. & COULANGE, C. 2007. Renal cancer and second cancer : critical review of the literature. *Progres En Urologie,* 17**,** 35-40.

NICE 2008. Prostate cancer: diagnosis and treatment CG28.

NICE 2014. Prostate cancer: diagnosis and treatment CG175. National Institute for Health and Care Excellence.

NICE 2019. Prostate cancer: diagnosis and management, NG131.

NILSSON, S., NORLEN, B. J. & WIDMARK, A. 2004. A systematic overview of radiation therapy effects in prostate cancer. *Acta Oncologica,* 43**,** 316-381.

NOVARA, G., FICARRA, V., MOCELLIN, S., AHLERING, T. E., CARROLL, P. R., GRAEFEN, M., GUAZZONI, G., MENON, M., PATEL, V. R., SHARIAT, S. F., TEWARI, A. K., VAN POPPEL, H., ZATTONI, F., MONTORSI, F., MOTTRIE, A., ROSEN, R. C. & WILSON, T. G. 2012. Systematic review and meta-analysis of studies reporting oncologic outcome after robot-assisted radical prostatectomy. *Eur Urol,* 62**,** 382-404.

OKEEFE, R. M., BALCI, O. & SMITH, E. P. 1987. VALIDATING EXPERT SYSTEM PERFORMANCE. *Ieee Expert-Intelligent Systems & Their Applications,* 2**,** 81-90.

OKEEFE, R. M. & OLEARY, D. E. 1993. EXPERT SYSTEM VERIFICATION AND VALIDATION - A SURVEY AND TUTORIAL. *Artificial Intelligence Review,* 7**,** 3-42.

OLIVER, S. E., DONOVAN, J. L., PETERS, T. J., FRANKEL, S., HAMDY, F. C. & NEAL, D. E. 2003. Recent trends in the use of radical prostatectomy in England: the epidemiology of diffusion. *BJU Int,* 91**,** 331-6; discussion 336.

OLSON, J. R. & RUETER, H. H. 1987. Extracting expertise from experts: Methods for knowledge acquisition. *Expert Systems,* 4**,** 152-168.

OLUFUNKE, O. O., CHARLES, U. O., CHARLES, A. K., ABRAHAM, A. & SNASEL, V. 2013. A FUZZY-MINING APPROACH FOR SOLVING RULE BASED EXPERT SYSTEM UNWIELDINESS IN MEDICAL DOMAIN. *Neural Network World,* 23**,** 435-450.

PANDEY, B. & MISHRA, R. B. 2009. Knowledge and intelligent computing system in medicine. *Computers in Biology and Medicine,* 39**,** 215-230.

PANTAZOPOULOS, D., KARAKITSOS, P., IOKIM-LIOSSI, A., POULIAKIS, A., BOTSOLI-STERGIOU, E. & DIMOPOULOS, C. 1998a. Back propagation neural network in the discrimination of benign from malignant lower urinary tract lesions. *Journal of Urology,* 159**,** 1619-1623.

PANTAZOPOULOS, D., KARAKITSOS, P., IOKIM-LIOSSI, A., POULIAKIS, A. & DIMOPOULOS, K. 1998b. Comparing neural networks in the discrimination of benign from malignant lower urinary tract lesions. *British Journal of Urology,* 81**,** 574-579.

PAPAGEORGIOU, E. I. 2012. Fuzzy cognitive map software tool for treatment management of uncomplicated urinary tract infection. *Computer Methods and Programs in Biomedicine,* 105**,** 233-245.

PAREKATTIL, S. J., FISHER, H. A. G. & KOGAN, B. A. 2003. Neural network using combined urine nuclear matrix protein-22, monocyte chemoattractant protein-1 and urinary intercellular adhesion molecule-1 to detect bladder cancer. *Journal of Urology,* 169**,** 917-920.

PAREKATTIL, S. J., GILL, I. S., CASTLE, E. P., BURGESS, S. V., WALLS, M. M., THOMAS, R., KUMAR, U., PURIFOY, J. A., NG, C. S., KANG, Y., FUCHS, G. J., WEISE, E. S., WINFIELD, H. N., LALLAS, C. & ANDREWS, P. E. 2005. Multi-institutional validation study of neural networks to predict duration of stay after laparoscopic radical/simple or partial nephrectomy. *Journal of Urology,* 174**,** 1380-1384.

PAYA, A. S., FERNANDEZ, D. R., MENDEZ, D. G., HERNANDEZ, C. A. M. & IEEE 2006. Development of an artificial neural network for helping to diagnose diseases in urology. *2006 1st Bio-Inspired Models of Network, Information and Computing Systems***,** 47-50.

PEREIRA, M., SCHAEFER, M. & B MARQUES, J. 2004. Remote expert system of support the prostate cancer diagnosis. *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference,* 5**,** 3412-5.

PETROVIC, S., MISHRA, N. & SUNDAR, S. 2011. A novel case based reasoning approach to radiotherapy planning. *Expert Systems with Applications,* 38**,** 10759-10769.

PETRUCCI, K., PETRUCCI, P., CANFIELD, K., MCCORMICK, K. A., KJERULFF, K. & PARKS, P. 1991. Evaluation of UNIS: Urological Nursing Information Systems. *Proceedings / the ... Annual Symposium on Computer Application [sic] in Medical Care. Symposium on Computer Applications in Medical Care*.

PILEPICH, M. V., KRALL, J. M., SAUSE, W. T., JOHNSON, R. J., RUSS, H. H., HANKS, G. E., PEREZ, C. A., ZINNINGER, M., MARTZ, K. L. & GARDNER, P. 1987. Correlation of radiotherapeutic parameters and treatment related morbidity in carcinoma of the prostate— analysis of rtog study 75-06. *International Journal of Radiation Oncology*Biology*Physics,* 13**,** 351-357.

PILEPICH, M. V., WINTER, K., LAWTON, C. A., KRISCH, R. E., WOLKOV, H. B., MOVSAS, B., HUG, E. B., ASBELL, S. O. & GRIGNON, D. 2005. Androgen suppression adjuvant to definitive radiotherapy in prostate carcinoma--long-term results of phase III RTOG 85-31. *Int J Radiat Oncol Biol Phys,* 61**,** 1285-90.

PORTER, C., O'DONNELL, C., CRAWFORD, E. D., GAMITO, E. J., ERREJON, A., GENEGA, E., SOTELO, T. & TEWARI, A. 2001. Artificial neural network model to predict biochemical failure after radical prostatectomy. *Molecular Urology,* 5**,** 159-162.

PORTER, C. R., GAMITO, E. J., CRAWFORD, E. D., BARTSCH, G., PRESTI, J. C., TEWARI, A. & O'DONNELL, C. 2005. Model to predict prostate biopsy outcome in large screening population with independent validation in referral setting. *Urology,* 65**,** 937-941.

POTTER, S. R., MILLER, M. C., MANGOLD, L. A., JONES, K. A., EPSTEIN, J. I., VELTRI, R. W. & PARTIN, A. W. 1999. Genetically engineered neural networks for predicting prostate cancer progression after radical prostatectomy. *Urology,* 54**,** 791-795.

POULAKIS, V., WITZSCH, U., DE VRIES, R., EMMERLICH, V., MEVES, M., ALTMANNSBERGER, H. M. & BECHT, E. 2004a. Preoperative neural network using combined magnetic resonance imaging variables, prostate-specific antigen, and Gleason score to predict positive surgical margins. *Urology,* 64**,** 516-521.

POULAKIS, V., WITZSCH, U., DE VRIES, R., EMMERLICH, V., MEVES, M., ALTMANNSBERGER, H. M. & BECHT, E. 2004b. Preoperative neural network using combined magnetic resonance imaging variables, prostate specific antigen and Gleason score to predict prostate cancer stage. *Journal of Urology,* 172**,** 1306-1310.

POULAKIS, V., WITZSCH, U., REMPLICK, J., SIHLER, S. & BECHT, E. 2002. Prediction of clearance of inferior caliceal calculi with extracorporeal shock wave lithotripsy. Using an artificial neural network analysis. *Urologe A,* 41**,** 583-595.

POWELL, C. R., DESAI, R. A., MAKHLOUF, A. A., SIGMAN, M., JAROW, J. P., ROSS, L. S. & NIEDERBERGER, C. S. 2007. Computational models for detection of endocrinopathy in subfertile males. *Int J Impot Res,* 20**,** 79-84.

PREECE, A. D., SHINGHAL, R. & BATAREKH, A. 1992. VERIFYING EXPERT SYSTEMS - A LOGICAL FRAMEWORK AND A PRACTICAL TOOL. *Expert Systems with Applications,* 5**,** 421-436.

PRICE, A. J., ALLEN, N. E., APPLEBY, P. N., CROWE, F. L., TRAVIS, R. C., TIPPER, S. J., OVERVAD, K., GRONBAEK, H., TJONNELAND, A., JOHNSEN, N. F., RINALDI, S., KAAKS, R., LUKANOVA, A., BOEING, H., ALEKSANDROVA, K., TRICHOPOULOU, A., TRICHOPOULOS, D., ANDARAKIS, G., PALLI, D., KROGH, V., TUMINO, R., SACERDOTE, C., BUENO-DE-MESQUITA, H. B., ARGUELLES, M. V., SANCHEZ, M. J., CHIRLAQUE, M. D., BARRICARTE, A., LARRANAGA, N., GONZALEZ, C. A., STATTIN, P., JOHANSSON, M., KHAW, K. T., WAREHAM, N., GUNTER, M., RIBOLI, E. & KEY, T. 2012. Insulin-like Growth Factor-I Concentration and Risk of Prostate Cancer: Results from the European Prospective Investigation into Cancer and Nutrition. *Cancer Epidemiology Biomarkers & Prevention,* 21**,** 1531-1541.

QUINN, M. & BABB, P. 2002. Patterns and trends in prostate cancer incidence, survival, prevalence and mortality. Part I: international comparisons. *Bju International,* 90**,** 162-173.

QURESHI, K. N., NAGUIB, R. N. G., HAMDY, F. C., NEAL, D. E. & MELLON, J. K. 2000. Neural network analysis of clinicopathological and molecular markers in bladder cancer. *Journal of Urology,* 163**,** 630-633.

RAMAKRISHNA, N. R. & DEWEESE, T. L. 2001. Radiation therapy as applied to prostate cancer - Clinical, technical, and biologic considerations. *Prostate Cancer***,** 387-413.

RAMASAMY, R., PADILLA, W. O., OSTERBERG, E. C., SRIVASTAVA, A., REIFSNYDER, J. E., NIEDERBERGER, C. & SCHLEGEL, P. N. 2013. A Comparison of Models for Predicting Sperm Retrieval Before Microdissection Testicular Sperm Extraction in Men with Nonobstructive Azoospermia. *The Journal of Urology,* 189**,** 638-642.

RAMSAY, C. R., ADEWUYI, T. E., GRAY, J., HISLOP, J., SHIRLEY, M. D., JAYAKODY, S., MACLENNAN, G., FRASER, C., MACLENNAN, S., BRAZZELLI, M., N'DOW, J., PICKARD, R., ROBERTSON, C., ROTHNIE, K., RUSHTON, S. P., VALE, L. & LAM, T. B. 2015. Ablative therapy for people with localised prostate cancer: a systematic review and economic evaluation. *Health Technol Assess,* 19**,** 1-490.

REGNIER-COUDERT, O., MCCALL, J., LOTHIAN, R., LAM, T., MCCLINTON, S. & N'DOW, J. 2012. Machine learning for improved pathological staging of prostate cancer: A performance comparison on a range of classifiers. *Artificial Intelligence in Medicine,* 55**,** 25-35.

REMZI, M., ANAGNOSTOU, T., RAVERY, V., ZLOTTA, A., STEPHAN, C., MARBERGER, M., DJAVAN, B. & ESOU 2003. An artificial neural network to predict the outcome of repeat prostate biopsies. *Urology,* 62**,** 456-460.

RENEHAN, A. G., ZWAHLEN, M., MINDER, C., O'DWYER, S. T., SHALET, S. M. & EGGER, M. 2004. Insulin-like growth factor (IGF)-I, IGF binding protein-3, and cancer risk: systematic review and meta-regression analysis. *Lancet,* 363**,** 1346-1353.

RODDA, S., TYLDESLEY, S., MORRIS, W. J., KEYES, M., HALPERIN, R., PAI, H., MCKENZIE, M., DUNCAN, G., MORTON, G., HAMM, J. & MURRAY, N. 2017. ASCENDE-RT: An Analysis of Treatment-Related Morbidity for a Randomized Trial Comparing a Low-Dose-Rate Brachytherapy Boost with a Dose-Escalated External Beam Boost for High- and Intermediate-Risk Prostate Cancer. *Int J Radiat Oncol Biol Phys,* 98**,** 286-295.

RODDAM, A. W., ALLEN, N. E., APPLEBY, P., KEY, T. J., FERRUCCI, L., CARTER, H. B., METTER, E. J., CHEN, C., WEISS, N. S., FITZPATRICK, A., HSING, A. W., LACEY, J. V., HELZLSOUER, K., RINALDI, S., RIBOLI, E., KAAKS, R., JANSSEN, J., WILDHAGEN, M. F., SCHRODER, F. H., PLATZ, E. A., POLLAK, M., GIOVANNUCCI, E., SCHAEFER, C., QUESENBERRY, C. P., VOGELMAN, J. H., SEVERI, G., ENGLISH, D. R., GILES, G. G., STATTIN, P., HALLMANS, G., JOHANSSON, M., CHAN, J. M., GANN, P., OLIVER, S. E., HOLLY, J. M., DONOVAN, J.,

MEYER, F., BAIRATI, I. & GALAN, P. 2008. Insulin-like growth factors, their binding proteins, and prostate cancer risk: Analysis of individual patient data from 12 prospective studies. *Annals of Internal Medicine,* 149**,** 461-W88.

RONCO, A. L. & FERNANDEZ, R. 1999. Improving ultrasonographic diagnosis of prostate cancer with neural networks. *Ultrasound in Medicine and Biology,* 25**,** 729-733.

ROSENBLATT, F. 1958. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review,* 65**,** 386-408.

ROWLANDS, M. A., GUNNELL, D., HARRIS, R., VATTEN, L. J., HOLLY, J. M. P. & MARTIN, R. M. 2009. Circulating insulin-like growth factor peptides and prostate cancer risk: A systematic review and meta-analysis. *International Journal of Cancer,* 124**,** 2416-2429.

RYAN, C. J., SMITH, M. R., FIZAZI, K., SAAD, F., MULDERS, P. F., STERNBERG, C. N., MILLER, K., LOGOTHETIS, C. J., SHORE, N. D., SMALL, E. J., CARLES, J., FLAIG, T. W., TAPLIN, M. E., HIGANO, C. S., DE SOUZA, P., DE BONO, J. S., GRIFFIN, T. W., DE PORRE, P., YU, M. K., PARK, Y. C., LI, J., KHEOH, T., NAINI, V., MOLINA, A. & RATHKOPF, D. E. 2015. Abiraterone acetate plus prednisone versus placebo plus prednisone in chemotherapy-naive men with metastatic castration-resistant prostate cancer (COU-AA-302): final overall survival analysis of a randomised, double-blind, placebo-controlled phase 3 study. *Lancet Oncol,* 16**,** 152-60.

SAMLI, M. M. & DOGAN, I. 2004. AN ARTIFICIAL NEURAL NETWORK FOR PREDICTING THE PRESENCE OF SPERMATOZOA IN THE TESTES OF MEN WITH NONOBSTRUCTIVE AZOOSPERMIA. *The Journal of Urology,* 171**,** 2354-2357.

SANDOW, J., VONRECHENBERG, W., JERZABEK, G. & STOLL, W. 1978. PITUITARY GONADOTROPIN INHIBITION BY A HIGHLY ACTIVE ANALOG OF LUTEINIZING-HORMONE-RELEASING HORMONE. *Fertility and Sterility,* 30**,** 205-209.

SARITAS, I., ALLAHVERDI, N. & SERT, I. U. 2003. A fuzzy expert system design for diagnosis of prostate cancer. *Proceedings of the 4th international conference conference on Computer systems and technologies: e-Learning.* Rousse, Bulgaria: ACM.

SARITAS, I., OZKAN, I. A. & SERT, I. U. 2010. Prognosis of prostate cancer by artificial neural networks. *Expert Systems with Applications,* 37**,** 6646-6650.

SCHALLY, A. V., COMARU-SCHALLY, A. M., PLONOWSKI, A., NAGY, A., HALMOS, G. & REKASI, Z. 2000. Peptide analogs in the therapy of prostate cancer. *Prostate,* 45**,** 158-166.

SCHER, H. I., FIZAZI, K., SAAD, F., TAPLIN, M.-E., STERNBERG, C. N., MILLER, K., DE WIT, R., MULDERS, P., CHI, K. N., SHORE, N. D., ARMSTRONG, A. J., FLAIG, T. W., FLÉCHON, A., MAINWARING, P., FLEMING, M., HAINSWORTH, J. D., HIRMAND, M., SELBY, B., SEELY, L. & DE BONO, J. S. 2012. Increased Survival with Enzalutamide in Prostate Cancer after Chemotherapy. *New England Journal of Medicine,* 367**,** 1187-1197.

SCHREIBER, G., WIELINGA, B., DEHOOG, R., AKKERMANS, H. & VANDEVELDE, W. 1994. COMMONKADS - A COMPREHENSIVE METHODOLOGY FOR KBS DEVELOPMENT. *Ieee Expert-Intelligent Systems & Their Applications,* 9**,** 28-37.

SCHUESSLER, W. W., SCHULAM, P. G., CLAYMAN, R. V. & KAVOUSSI, L. R. 1997. Laparoscopic radical prostatectomy: initial short-term experience. *Urology,* 50**,** 854-7.

SEÇKINER, I., SEÇKINER, S. U., BAYRAK, Ö. & ERTURHAN, S. 2011. Use of artificial neural networks in the management of antenatally diagnosed ureteropelvic junction obstruction. *Canadian Urological Association Journal,* 5**,** E152.

SEIDENFELD, J., SAMSON, D. J., ARONSON, N., ALBERTSON, P. C., BAYOUMI, A. M., BENNETT, C., BROWN, A., GARBER, A., GERE, M., HASSELBLAD, V., WILT, T. & ZIEGLER, K. 1999. Relative effectiveness and cost-effectiveness of methods of androgen suppression in the treatment of advanced prostate cancer. *Evidence report/technology assessment (Summary)***,** i-x, 1-246, I1-36, passim.

SEIDENFELD, J., SAMSON, D. J., HASSELBLAD, V., ARONSON, N., ALBERTSEN, P. C., BENNETT, C. L. & WILT, T. J. 2000. Single-therapy androgen suppression in men with advanced

prostate cancer: A systematic review and meta-analysis. *Annals of Internal Medicine,* 132**,** 566-577.

SEKER, H., ODETAYO, M. O., PETROVIC, D. & NAGUIB, R. N. G. 2003. A fuzzy logic based-method for prognostic decision making in breast and prostate cancers. *Ieee Transactions on Information Technology in Biomedicine,* 7**,** 114-122.

SELVADURAI, E. D., SINGHERA, M., THOMAS, K., MOHAMMED, K., WOODE-AMISSAH, R., HORWICH, A., HUDDART, R. A., DEARNALEY, D. P. & PARKER, C. C. 2013. Medium-term Outcomes of Active Surveillance for Localised Prostate Cancer. *European Urology,* 64**,** 981-987.

SERATI, M., SALVATORE, S., SIESTO, G., CATTONI, E., BRAGA, A., SORICE, P., CROMI, A., GHEZZI, F. & BOLIS, P. 2011. Urinary Symptoms and Urodynamic Findings in Women with Pelvic Organ Prolapse: Is There a Correlation? Results of an Artificial Neural Network Analysis. *European Urology,* 60**,** 253-260.

SERRANO-DURBA, A., SERRANO, A. J., MAGDALENA, J. R., MARTIN, J. D., SORIA, E., DOMINGUEZ, C., ESTORNELL, F. & GARCIA-IBARRA, F. 2004. The use of neural networks for predicting the result of endoscopic treatment for vesico-ureteric reflux. *Bju International,* 94**,** 120-122.

SHORTLIFFE, E. H., AXLINE, S. G., BUCHANAN, B. G., CHAVEZPARDO, R., DAVIS, R., RHAME, F., SCOTT, C., VANMELLE, W. & COHEN, S. N. 1975. COMPUTER AS A CONSULTANT FOR SELECTION OF ANTIMICROBIAL THERAPY FOR PATIENTS WITH BACTEREMIA. *Clinical Research,* 23**,** A385-A385.

SPECIALISED COMMISSIONING TEAM, N. E. 2015. Clinical Commissioning

Policy: Robotic-Assisted

Surgical Procedures for

Prostate Cancer, Reference: NHSEngland B14/P/a. *In:* SPECIALISED COMMISSIONING TEAM, N. E. (ed.).

STAMEY, T. A., BARNHILL, S. D., ZHANG, Z., MADYASTHA, K. R., PRESTIGIACOMO, A. F., JONES, K. & CHAN, D. 1996. Effectiveness of prostasure in detecting prostate cancer (PCa) and benign prostatic hyperplasia (BPH) in men age 50 and older. *Journal of Urology,* 155**,** 436A.

STEPHAN, C., CAMMANN, H., BENDER, M., MILLER, K., LEIN, M., JUNG, K. & MEYER, H. A. 2010. Internal validation of an artificial neural network for prostate biopsy outcome. *International Journal of Urology,* 17**,** 62-68.

STEPHAN, C., CAMMANN, H., MEYER, H. A., MULLER, C., DEGER, S., LEIN, M. & JUNG, K. 2008. An artificial neural network for five different assay systems of prostate-specific antigen in prostate cancer diagnostics. *Bju International,* 102**,** 799-805.

STEPHAN, C., CAMMANN, H., SEMJONOW, A., DIAMANDIS, E. P., WYMENGA, L. F. A., LEIN, M., SINHA, P., LOENING, S. A. & JUNG, K. 2002. Multicenter evaluation of an artificial neural network to increase the prostate cancer detection rate and reduce unnecessary biopsies. *Clinical Chemistry,* 48**,** 1279-1287.

STEPHAN, C., JUNG, K., SOOSAIPILLAI, A., YOUSEF, G. M., CAMMANN, H., MEYER, H., XU, C. L. & DIAMANDIS, E. P. 2005. Clinical utility of human glandular kallikrein 2 within a neural network for prostate cancer detection. *Bju International,* 96**,** 521-527.

STEPHAN, C., KAHRS, A. M., CAMMANN, H., LEIN, M., SCHRADER, M., DEGER, S., MILLER, K. & JUNG, K. 2009. A -2 proPSA-Based Artificial Neural Network Significantly Improves Differentiation Between Prostate Cancer and Benign Prostatic Diseases. *Prostate,* 69**,** 198-207.

STEPHAN, C., MEYER, H. A., CAMMANN, H., NAKAMURA, T., DIAMANDIS, E. P. & JUNG, K. 2006a. Improved prostate cancer detection with a human kallikrein 11 and percentage free PSA-based artificial neural network. *Biological Chemistry,* 387**,** 801-805.

STEPHAN, C., MEYER, H. A., KWIATKOWSKI, M., RECKER, F., CAMMANN, H., LOENING, S. A., JUNG, K. & LEIN, M. 2006b. A (-5,-7) ProPSA based artificial neural network to detect prostate cancer. *European Urology,* 50**,** 1014-1020.

STEPHAN, C., VINCENDEAU, S., HOULGATTE, A., CAMMANN, H., JUNG, K. & SEMJONOW, A. 2013. Multicenter Evaluation of -2 Proprostate-Specific Antigen and the Prostate Health Index for Detecting Prostate Cancer. *Clinical Chemistry,* 59**,** 306-314.

STEPHAN, C., XU, C. L., FINNE, P., CAMMANN, H., MEYER, H. A., LEIN, M., JUNG, K. & STENMAN, U. H. 2007. Networks for prostate biopsy indication in two different patient Populations comparison of two different artificial neural. *Urology,* 70**,** 596-601.

STEPHENSON, A. J., KATTAN, M. W., EASTHAM, J. A., DOTAN, Z. A., BIANCO, F. J., LILJA, H. & SCARDINO, P. T. 2006. Defining biochemical recurrence of prostate cancer after radical prostatectomy: A proposal for standardized definition. *Journal of Clinical Oncology,* 24**,** 3973-3978.

SUCEVIC, D. & ILIC, I. 1991. UNCERTAIN KNOWLEDGE PROCESSING IN UROLOGY DIAGNOSTIC PROBLEMS BASED EXPERT SYSTEM. *6th Mediterranean Electrotechnical Conference, Proceedings Vols 1 and 2***,** 741-743.

SUN, C. C. & CHANG, P. 2006. Prediction of unexpected emergency room visit after extracorporeal shock wave lithotripsy for urolithiasis - an application of artificial neural network in hospital information system. *AMIA Annu Symp Proc***,** 1113.

SWEENEY, C. J., CHEN, Y. H., CARDUCCI, M., LIU, G., JARRARD, D. F., EISENBERGER, M., WONG, Y. N., HAHN, N., KOHLI, M., COONEY, M. M., DREICER, R., VOGELZANG, N. J., PICUS, J., SHEVRIN, D., HUSSAIN, M., GARCIA, J. A. & DIPAOLA, R. S. 2015. Chemohormonal Therapy in Metastatic Hormone-Sensitive Prostate Cancer. *N Engl J Med,* 373**,** 737-46.

TANTHANUCH, M. & TANTHANUCH, S. 2004. Prediction of upper urinary tract calculi using an artificial neural network. *J Med Assoc Thai,* 87**,** 515-8.

TEWARI, A., ISSA, M., EL-GALLEY, R., STRICKER, H., PEABODY, J., POW-SANG, J., SHUKLA, A., WAJSMAN, Z., RUBIN, M., WEI, J., MONTIE, J., DEMERS, R., JOHNSON, C. C., LAMERATO, L., DIVINE, G. W., CRAWFORD, E. D., GAMITO, E. J., FARAH, R., NARAYAN, P., CARLSON, G. & MENON, M. 2001. Genetic adaptive neural network to predict biochemical failure after radical prostatectomy: a multi-institutional study. *Mol Urol,* 5**,** 163-9.

TEWARI, A. & NARAYAN, P. 1998. Novel staging tool for localized prostate cancer: A pilot study using genetic adaptive neural networks. *Journal of Urology,* 160**,** 430-436.

THOMPSON, I. M., TANGEN, C. M., PARADELO, J., LUCIA, M. S., MILLER, G., TROYER, D., MESSING, E., FORMAN, J., CHIN, J., SWANSON, G., CANBY-HAGINO, E. & CRAWFORD, E. D. 2009. Adjuvant Radiotherapy for Pathological T3N0M0 Prostate Cancer Significantly Reduces Risk of Metastases and Improves Survival: Long-Term Followup of a Randomized Clinical Trial. *The Journal of Urology,* 181**,** 956-962.

TOLIS, G., ACKMAN, D., STELLOS, A., MEHTA, A., LABRIE, F., FAZEKAS, A. T. A., COMARUSCHALLY, A. M. & SCHALLY, A. V. 1982. TUMOR-GROWTH INHIBITION IN PATIENTS WITH PROSTATIC-CARCINOMA TREATED WITH LUTEINIZING-HORMONE-RELEASING HORMONE AGONISTS. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences,* 79**,** 1658-1662.

TORSHIZI, A. D., ZARANDI, M. H. F., TORSHIZI, G. D. & EGHBALI, K. 2014. A hybrid fuzzy-ontology based intelligent system to determine level of severity and treatment recommendation for Benign Prostatic Hyperplasia. *Computer Methods and Programs in Biomedicine,* 113**,** 301-313.

TSAO, C. W., LIU, C. Y., CHA, T. L., WU, S. T., SUN, G. H., YU, D. S., CHEN, H. I., CHANG, S. Y., CHEN, S. C. & HSU, C. Y. 2014. Artificial neural network for predicting pathological stage

of clinically localized prostate cancer in a Taiwanese population. *Journal of the Chinese Medical Association,* 77**,** 513-518.

TURBAN, E. 1990. *Decision Support and Expert Systems: Management Support Systems*, Prentice Hall PTR.

TURING, A. M. 1950. Computing machinery and intelligence. *Mind,* 59**,** 433-460.

TYRRELL, C. J., PAYNE, H., SEE, W. A., MCLEOD, D. G., WIRTH, M. P., IVERSEN, P., ARMSTRONG, J., MORRIS, C. & CASODEX EARLY PROSTATE CANC, T. 2005. Bicalutamide ('Casodex') 150 mg as adjuvant to radiotherapy in patients with localised or locally advanced prostate cancer: Results from the randomised Early Prostate Cancer Programme. *Radiotherapy and Oncology,* 76**,** 4-10.

UK, C. R. 2011. *Cancer Stats report - Bladder Cancer* [Online]. Available: http://www.cancerresearchuk.org/cancer-info/cancerstats/incidence/commoncancers/.

VALE, C. L., BURDETT, S., RYDZEWSKA, L. H. M., ALBIGES, L., CLARKE, N. W., FISHER, D., FIZAZI, K., GRAVIS, G., JAMES, N. D., MASON, M. D., PARMAR, M. K. B., SWEENEY, C. J., SYDES, M. R., TOMBAL, B. & TIERNEY, J. F. 2016. Addition of docetaxel or bisphosphonates to standard of care in men with localised or metastatic, hormone-sensitive prostate cancer: a systematic review and meta-analyses of aggregate data. *The Lancet Oncology,* 17**,** 243-256.

VANHAECHT, K., DE WITTE, K., DEPREITERE, R. & SERMEUS, W. 2006. Clinical pathway audit tools: a systematic review. *Journal of nursing management,* 14**,** 529-37.

VELTRI, R. W., CHAUDHARI, M., MILLER, M. C., POOLE, E. C., O'DOWD, G. J. & PARTIN, A. W. 2002. Comparison of logistic regression and neural net modeling for prediction of prostate cancer pathologic stage. *Clin Chem,* 48**,** 1828-34.

VILCHEZMARTINEZ, J. A., PEDROZA, E., ARIMURA, A. & SCHALLY, A. V. 1979. PARADOXICAL EFFECTS OF D-TRP6-LUTEINIZING HORMONE-RELEASING HORMONE ON THE HYPOTHALAMIC-PITUITARY-GONADAL AXIS IN IMMATURE FEMALE RATS. *Fertility and Sterility,* 31**,** 677-682.

VOLMER, M., WOLTHERS, B. G., METTING, H. J., DEHAAN, T. H. Y., COENEGRACHT, P. M. J. & VANDERSLIK, W. 1994. ARTIFICIAL NEURAL-NETWORK PREDICTIONS OF URINARY CALCULUS COMPOSITIONS ANALYZED WITH INFRARED-SPECTROSCOPY. *Clinical Chemistry,* 40**,** 1692-1697.

VON DER MAASE, H., SENGELOV, L., ROBERTS, J. T., RICCI, S., DOGLIOTTI, L., OLIVER, T., MOORE, M. J., ZIMMERMANN, A. & ARNING, M. 2005. Long-term-survival results of a randomized trial comparing gemcitabine plus cisplatin, with methotrexate, vinblastine, doxorubicin, plus cisplatin in patients with bladder cancer (Retracted article. See vol. 16, pg. 1481, 2011). *Journal of Clinical Oncology,* 23**,** 4602-4608.

VUKICEVIC, A. M., JOVICIC, G. R., STOJADINOVIC, M. M., PRELEVIC, R. I. & FILIPOVIC, N. D. 2014. Evolutionary assembled neural networks for making medical decisions with minimal regret: Application for predicting advanced bladder cancer outcome. *Expert Systems with Applications,* 41**,** 8092-8100.

W. HERR, H., DOTAN, Z., DONAT, S. M. & BAJORIN, D. F. 2007. Defining optimal therapy for muscle invasive bladder cancer. *Journal of Urology,* 177**,** 437-443.

WADIE, B. S., BADAWI, A. M., ABDELWAHED, M. & ELEMABAY, S. M. 2006. Application of artificial neural network in prediction of bladder outlet obstruction: a model based on objective, noninvasive parameters. *Urology,* 68**,** 1211-4.

WADIE, B. S., BADAWI, A. M. & GHONEIM, M. A. 2001. THE RELATIONSHIP OF THE INTERNATIONAL PROSTATE SYMPTOM SCORE AND OBJECTIVE PARAMETERS FOR DIAGNOSING BLADDER OUTLET OBSTRUCTION. PART II: THE POTENTIAL USEFULNESS OF ARTIFICIAL NEURAL NETWORKS. *The Journal of Urology,* 165**,** 35-37.

WALSH, P. C., LEPOR, H. & EGGLESTON, J. C. 1983. Radical prostatectomy with preservation of sexual function: Anatomical and pathological considerations. *The Prostate,* 4**,** 473-485.

WANG, G., LAM, K.-M., DENG, Z. & CHOI, K.-S. 2015. Prediction of mortality after radical cystectomy for bladder cancer by machine learning techniques. *Computers in biology and medicine,* 63**,** 124-132.

WATSON, E. K., O'BRIEN, R., CAMPBELL, C., WELLER, D., NEAL, R. D., WILKINSON, C., ROSE, P. W. & PROSTATE CANC FOLLOW-UP STUDY, G. 2011. Views of health professionals on the role of primary care in the follow-up of men with prostate cancer. *Family Practice,* 28**,** 647-654.

WELLS, D. M. & NIEDERER, J. 1998. A medical expert system approach using artificial neural networks for standardized treatment planning. *International Journal of Radiation Oncology Biology Physics,* 41**,** 173-182.

WERBOS, P. J. 1974. *Beyond regression: New tools for prediction and analysis in the behavioral sciences.* Ph.D, Harward.

WHITMORE, W. F., GRABSTAL.H & HILARIS, B. 1972. RETROPUBIC IMPLANTATION OF IODINE 125 IN TREATMENT OF PROSTATIC CANCER. *Journal of Urology,* 108**,** 918-&.

WILLIAMS, N., HUGHES, L. J., TURNER, E. L., DONOVAN, J. L., HAMDY, F. C., NEAL, D. E., MARTIN, R. M. & METCALFE, C. 2011. Prostate-specific antigen testing rates remain low in UK general practice: a cross-sectional study in six English cities. *Bju International,* 108**,** 1402-1408.

XIAO, D., ZHANG, G., LIU, Y., YANG, Z., ZHANG, X., LI, L., JIAO, C. & LU, H. 2016. 3D detection and extraction of bladder tumors via MR virtual cystoscopy. *International Journal of Computer Assisted Radiology and Surgery,* 11**,** 89-97.

YANG, H. L. 1995. INFORMATION KNOWLEDGE ACQUISITION METHODS FOR DECISION-SUPPORT SYSTEMS AND EXPERT-SYSTEMS. *Information Processing & Management,* 31**,** 47-58.

YOUNG, H. H. 1903. Conservative perineal prostatectomy. A presentation of new instruments and technic. *Journal of the American Medical Association,* 41**,** 999-1009.

YU, V. L., FAGAN, L. M., WRAITH, S. M., CLANCEY, W. J., SCOTT, A. C., HANNIGAN, J., BLUM, R. L., BUCHANAN, B. G. & COHEN, S. N. 1979. ANTI-MICROBIAL SELECTION BY A COMPUTER - BLINDED EVALUATION BY INFECTIOUS-DISEASES EXPERTS. *Jama-Journal of the American Medical Association,* 242**,** 1279-1282.

YUKSEL, S., DIZMAN, T., YILDIZDAN, G. & SERT, U. 2013. Application of soft sets to diagnose the prostate cancer risk. *Journal of Inequalities and Applications*.

ZADEH, L. A. 1965. Fuzzy sets. *Information and Control,* 8**,** 338-353.

ZAPATERO, A., GUERRERO, A., MALDONADO, X., ALVAREZ, A., GONZALEZ SAN SEGUNDO, C., CABEZA RODRIGUEZ, M. A., MACIAS, V., PEDRO OLIVE, A., CASAS, F., BOLADERAS, A., DE VIDALES, C. M., VAZQUEZ DE LA TORRE, M. L., VILLA, S., PEREZ DE LA HAZA, A. & CALVO, F. A. 2015. High-dose radiotherapy with short-term or long-term androgen deprivation in localised prostate cancer (DART01/05 GICOR): a randomised, controlled, phase 3 trial. *Lancet Oncol,* 16**,** 320-7.

ZELEFSKY, M. J., LEVIN, E. J., HUNT, M., YAMADA, Y., SHIPPY, A. M., JACKSON, A. & AMOLS, H. I. 2008. Incidence of Late Rectal and Urinary Toxicities After Three-Dimensional Conformal Radiotherapy and Intensity-Modulated Radiotherapy for Localized Prostate Cancer. *International Journal of Radiation Oncology*Biology*Physics,* 70**,** 1124-1129.

ZLOTTA, A. R., REMZI, M., SNOW, P. B., SCHULMAN, C. C., MARBERGER, M. & DJAVAN, B. 2003. An artificial neural network for prostate cancer staging when serum prostate specific antigen is 10 NG./ML. or less. *Journal of Urology,* 169**,** 1724-1728.

ZUCCOLO, L., HARRIS, R., GUNNELL, D., OLIVER, S., LANE, J. A., DAVIS, M., DONOVAN, J., NEAL, D., HAMDY, F., BEYNON, R., SAVOVIC, J. & MARTIN, R. M. 2008. Height and prostate cancer risk: A large nested case-control study (ProtecT) and meta-analysis. *Cancer Epidemiology Biomarkers & Prevention,* 17**,** 2325-2336.

Appendices

Appendix 1: Protocol of the clinical investigation plan

CLINICAL INVESTIGATION PLAN

A Clinical Evaluation of a Clinical Decision Support Software (CDSS) System for Stable Prostate Cancer (SPC)

Study No. RD-5103-006-09
iQudos Study No. CI2010-02

September 2010
Version 1.0
Author: Sue Hagan
Date of Approval: 15th September 2010
Date of Approval from R&D Office: 7th December 2010

DOCUMENT VERSION HISTORY		
VERSION	DATE	SUMMARY OF CHANGES
001	15th September 2010	

STUDY PERSONNEL

Principal Clinical Investigator	Stephen Thomas MBBS FRCS(Urol) Consultant Urologist Department of Urology Royal Derby Hospital, Uttoxeter Road, Derby DE22 3NE 01332 347141
Investigators	Amjad Peracha MRCS, LRCP, MBBS Consultant Urologist Jaskarn Rai BPharm, MBChB, MRCS Urology Registrar Department of Urology Royal Derby Hospital, Uttoxeter Road, Derby DE22 3NE 01332 347141
Sponsor	Teresa Grieve BSc (Hons), MBA, PhD Assistant Director of Research and Development Royal Derby Hospital, Uttoxeter Road, Derby DE22 3NE 01332 347141
Monitors	Sue Hagan BSc (Hons), MRPharmS, PhD Head of Clinical and Regulatory Affairs MS3 Ltd aka iQudos, Mansfield i-centre, Oakham Business Park, Hamilton Way, Mansfield, Nottinghamshire, NG18 5BR 07721 088285 DHFT Monitor Acting Research Governance Manager Royal Derby Hospital, Uttoxeter Road, Derby DE22 3NE 01332 347141
Medical Expert	Alastair McCance MA DM FRCP Clinical Director of Medicine Royal Derby Hospital, Uttoxeter Road, Derby DE22 3NE 01332 347141
Medical Statistician	Apostolos Fakis Medical Statistician Royal Derby Hospital, Uttoxeter Road, Derby DE22 3NE 01332 347141
Clinical investigation site	Royal Derby Hospital Uttoxeter Road, Derby DE22 3NE

Table of Contents

STUDY PERSONNEL 2
OVERALL SUMMARY OF THE CLINICAL INVESTIGATION 4
FLOW CHART SUMMARY OF CLINICAL INVESTIGATION 6
1. IDENTIFICATION AND DESCRIPTION OF THE MEDICAL DEVICE 9
2. PRELIMINARY INVESTIGATIONS AND JUSTIFICATION FOR THE STUDY 10
A. LITERATURE REVIEW 10
B. SUMMARY OF PRECLINICAL TESTING 11
C. PREVIOUS CLINICAL EXPERIENCE 12
D. DEVICE RISK ANALYSIS AND RISK ASSESSMENT 13
3. OBJECTIVES OF THE CLINICAL INVESTIGATION 15
4. DESIGN OF THE CLINICAL INVESTIGATION 16
A. METHODS 17
5. STATISTICAL CONSIDERATIONS 20
6. DEVIATIONS FROM THE CLINICAL INVESTIGATION PLAN 21
7. AMENDMENTS TO THE CLINICAL INVESTIGATION PLAN 22
8. ADVERSE EVENTS AND ADVERSE DEVICE EFFECTS 22
9. EARLY TERMINATION OR SUSPENSION OF THE INVESTIGATION 22
10. PUBLICATION POLICY 22
11. CASE REPORT FORMS (CRF) 22
12. MONITORING ARRANGEMENTS 22
13. DATA AND QUALITY MANAGEMENT 23
14. REFERENCES 25
15. APPENDICES 26
APPENDIX ONE 27
APPENDIX TWO 43
APPENDIX THREE 44

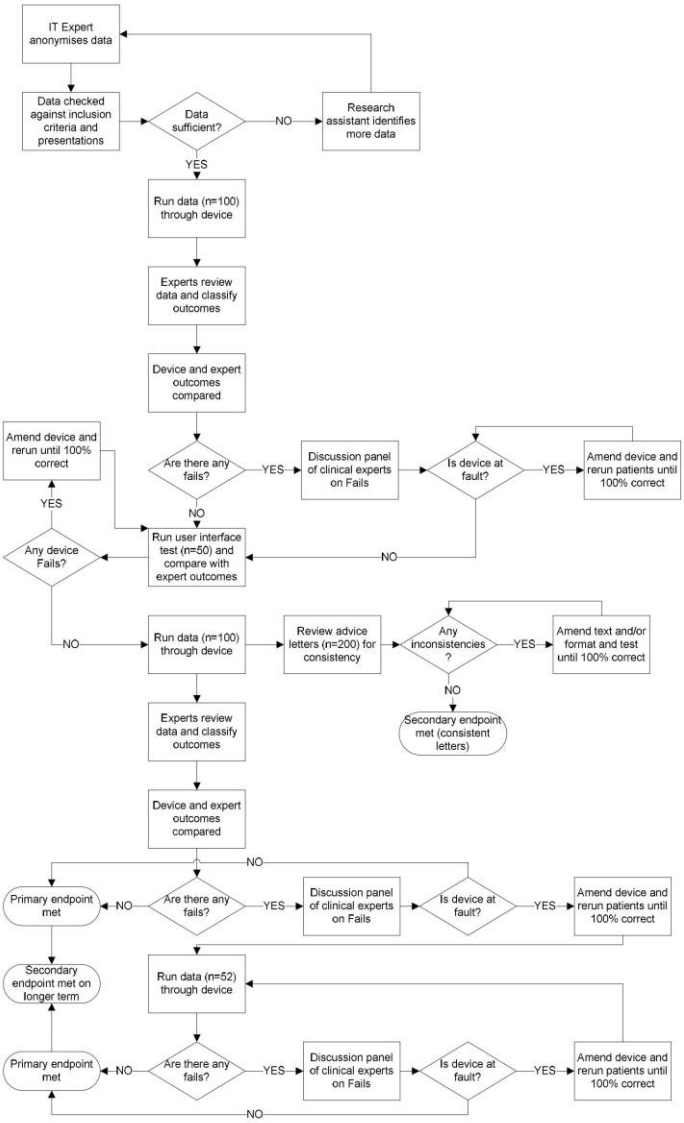
THIS REPORT IS THE PROPERTY OF DHFT. IT MUST BE KEPT IN A CONFIDENTIAL MANNER. NO UNPUBLISHED INFORMATION CONTAINED HEREIN WILL BE PUBLISHED OR DISCLOSED WITHOUT THE PRIOR WRITTEN APPROVAL OF DHFT.

Overall Summary of the Clinical Investigation

Title	A Clinical Evaluation of a Clinical Decision Support Software (CDSS) System for Stable Prostate Cancer (SPC).
Short Title	Stable Prostate Cancer CDSS v3.0 clinical investigation v1.0
Plan Version Number and Date	Version 1.0, August 2010.
Methodology	Clinical investigation of clinical decision support software using an anonymised retrospective database of clinical data. Clinical experts will review the subject data and make management recommendations that will then be compared to the recommendations made by the SPC CDSS. The clinical experts will review any discrepancies as a panel and the device corrected as necessary.
Study Duration	Up to 12 weeks. The overall duration of the study including report will be up to 18 weeks.
Study Centres	Single centre. Royal Derby Hospital, Derby.
Objectives	SPC CDSS system version 3.0 demonstrated to produce recommendations that are clinically accurate and comply with the NICE, BAUS and EAU guidelines without deviation.
Number of Subjects/Patients	Anonymised data from 200 patients, followed by sets of 52 patients.
Main Inclusion Criteria	Records of anonymised male patients diagnosed with stable prostate cancer that include a full record of all of the data required by the software.

Statistical Methodology and Analysis	The outcome from the investigators will be compared to that from the device for estimating their inter rater agreement using the Kappa coefficient (95% Confidence Interval). The classification of the patients by three different investigators based on the device letter will be compared using the Kappa coefficient (95% Confidence Interval). The percentage of agreement between the investigators and the device will be recorded after each subset of patients. To be correct, the management of all patients by the device must be clinically accurate and reasonable and comply with the applicable guidelines as judged by the three investigators. This would constitute a Pass. If there is not agreement with the device outcome in the management of all patients after discussion of the investigators as a panel, then the device is not managing patients correctly and this would constitute a Fail. The management should be correct in 100% of cases. The sensitivity, specificity, positive and negative predicted values (95% Confidence Interval) will be estimated for the accuracy of the software decision against the correct clinical decision.

Flow Chart Summary of Clinical Investigation



SIGNATURES

THIS PROTOCOL IS APPROVED BY:

SPONSOR, TERESA GRIEVE
SIGN DATE ☐☐ / ☐☐☐ / ☐☐
TERESA GRIEVE, ASSISTANT DIRECTOR OF R&D, DHFT

AUTHOR AND CLINICAL INVESTIGATION MONITOR/PROJECT MANAGER
SIGN DATE ☐☐ / ☐☐☐ / ☐☐
SUE HAGAN, HEAD OF CLINICAL AND REGULATORY AFFAIRS, IQUDOS

MEDICAL STATISTICIAN
SIGN DATE ☐☐ / ☐☐☐ / ☐☐
APOSTOLOS FAKIS, MEDICAL STATISTICIAN, DHFT

THIS INVESTIGATION WILL BE CONDUCTED IN COMPLIANCE WITH THIS CLINICAL INVESTIGATION PLAN, THE RELEVANT PARTS OF BS EN ISO 14155 PARTS 1 AND 2:2009 AND CURRENT MHRA, EC AND OTHER APPLICABLE REGULATORY REQUIREMENTS, INCLUDING THE COMPANY GOVERNANCE REQUIREMENTS.

I HAVE READ THE PROTOCOL AND I AGREE TO CONDUCT THE INVESTIGATION IN COMPLIANCE WITH THE INFORMATION CONTAINED IN THIS CLINICAL INVESTIGATION PLAN, THE RELEVANT PARTS OF BS EN ISO 14155 PARTS 1 AND 2:2009 AND CURRENT MHRA, EC AND OTHER APPLICABLE REGULATORY REQUIREMENTS, INCLUDING THE COMPANY GOVERNANCE REQUIREMENTS.

PRINCIPAL CLINICAL INVESTIGATOR
SIGN DATE ☐☐ / ☐☐☐ / ☐☐
STEPHEN THOMAS, CONSULTANT UROLOGIST, DHFT

1. Identification and Description of the Medical Device

This document outlines the plan for a clinical investigation of version 3.0 of a clinical decision support software system used to advise general practitioners in the management of patients with stable prostate cancer (SPC). The proprietary name of the device is the Clinical Decision Support Software (CDSS) web application with SPC module and is part of a clinical service provided to general practitioners (GPs) by registered trained nurses working for the Company (MS3 Ltd aka iQudos). The device that is the subject of this plan is known as SPC CDSS version 3.0. The manufacturer of the device is MS3 Ltd aka iQudos, Mansfield i-centre, Oakham Business Park, Hamilton Way, Mansfield, Nottinghamshire, NG18 5BR, UK; telephone 0845 2581469 and email admin@iQudos.com.

The intended purpose of the SPC CDSS v3.0 device is to advise GPs in the care of male patients who have been previously diagnosed with stable prostate cancer. It is only suitable for this purpose. Currently men with this condition are often seen in a hospital setting. This software enables men to be seen in the community setting rather than at a hospital. Patients who have acutely diagnosed prostate cancer or who do not have a previously decided management pathway are specifically excluded.

The device comprises the SPC clinical decision support software (CDSS) version 3.0 and is based on a secure remote server held within the NHS 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. As such, there are no storage and handling requirements and the device has no contact with tissues or body fluids. The device uses standard computer software and databases and is browser-based, with the user interface delivered as online forms to be completed only by the nurse user who is a company employee. This nurse is trained to use the software, either by one-to-one training or by interactive CD ROM, and to understand its limitations. Retraining will be provided on an annual basis and must be current. A secure access code will be required to access the software on the N3 server. The software can only be used over the N3 network from the secured server hosted within an NHS data centre.

The software consists of a database and a rule based decision engine. The software is developed in a Microsoft development environment using C# .NET Framework 3.5, ASP.NET and client scripting languages such as javascript. The database contains the patient records and data collected during the operation of the service. The algorithm for the software is derived from accepted clinical guidance for SPC from NICE (National Institute for Health and Clinical Excellence) (1), BAUS (British Association of Urological Surgeons) (2) and EAU (European Association of Urology) (3). These collectively are the applicable guidelines on which the software is based.

IQudos aim to provide a clinical service to General Practitioners (GPs) that includes the use of the SPC CDSS v3.0 device. The CDSS will be used by registered nurses and is designed to support the nurse in providing a local service to patients in primary care rather than a hospital setting. A Consultant Urological Surgeon will make the management decision in a hospital. This may be a decision to remove the patient's prostate, give radiotherapy or place the patient on surveillance. After the acute phase of treatment patients are suitable for follow up. Although the majority of men will remain well in the long term, it is important that those who are developing problems or showing signs of relapse are rapidly identified. Some of the signs of relapse are subtle leading to a resistance among GPs and Consultants for these patients to be managed in the community as suggested by the NICE. The iQudos nurses will hold clinics locally or remotely. Data from the patient's history and tests are entered into the software database via a data entry form on the web browser. These data are then interpreted using an algorithm based on the applicable guidelines; this provides early identification of patients developing problems. A paper or electronic advice letter is sent to the patient's GP advising on the monitoring performed, the resulting reasonable and safe management of the patient and detailing any further tests and follow up required. For patients suitable for management in primary care, the CDSS may recommend follow up appointments for further monitoring. For patients not suitable for management in primary care, referral to secondary care will be recommended. The GP maintains control of the treatment and thus clinical responsibility for the patient. No precautions are required for the device after use in the service.

2. Preliminary investigations and justification for the study

a. Literature review

The literature review conducted to support this clinical investigation is described in document 494 "Literature Review Protocol for the Stable Prostate Cancer Clinical Decision Support Software System" and document 496 "Literature Review Report for the Stable Prostate Cancer Clinical Decision Support Software System". These documents are appended in Appendix Four and Appendix Five and contain a list of all the literature reviewed.

There are two publications that relate to the LUTS CDSS device version 0.1 (4, 5; documents 387 and 386 respectively), a similar device used for a different indication (lower urinary tract symptoms (LUTS) rather than SPC). These papers summarise the early clinical data with the prototype device (v0.1). There are two other publications that discuss decision support systems to recommend management for LUTS, one computerised and one non-computerised. The computerised system (6) used an MS Office-based system to analyse patient history and data based on agreed rules in a prospective study in secondary care and then produced an advice letter to the GP. The outcomes were compared to that of

a consultant urologist and any differences in opinion discussed with the consultant. The non-computerised system (7) was a prospective study using a checklist for GPs for the management of LUTS in primary care based on validated formalised international expert panel judgements. The efficiency of the checklists was compared to the outcomes from an expert panel of urologists. These two studies follow an equivalent route for clinical investigation to the study described in this plan in that outcomes from the decision support tools are compared to that of clinical experts, justifying the design presented here.

A group in France computerised the guidelines for diagnosis, treatment and follow-up of patients with prostate and breast cancer for the Lorraine region, as both a website and as "Kasimir" software (8). They compared the use of paper guidelines with computerised guidelines and found that for breast cancer, the computerised guidelines were significantly more used whereas for prostate cancer there was no significant difference. A Canadian group are also studying computerising prostate cancer clinical care pathways, including management (9). However there is little detail of the clinical management proposed by the software or any associated studies, with the paper focussing more on the development of the software itself. Abbod et al (10) reviewed artificial intelligence systems for urological cancers, including prostate. However, none of the systems discussed included indications relevant to the SPC CDSS device.

No other papers were found with SPC and decision support systems using any alternative investigation methods. Like the study presented in this plan, neither published study presented any further controls other than comparison with clinical experts and both aimed for complete agreement between the clinical experts and the decision support systems.

b. Summary of preclinical testing

This device is based on new technology that is designed to form the basis of several similar products for different disease domains, with the SPC CDSS device being the second product in this group, following the LUTS CDSS device. The device consists of a software application, which is used to capture key elements of data during a nurse-led consultation and then process these data through a 'Decision Engine' component to arrive at accurate medical outcomes in the form of advice.

The device testing can be divided into two discrete elements. The first is functional testing of the software application user interface (UI), which ensures that all measures possible are taken to eliminate bugs in the GUI (Graphical User Interface) of the software. Several rounds of functional testing have been performed and the issues recorded in bug tracking systems to stabilise the data capture elements of the system. Several issues (n=59) were identified from this testing, leading to device amendments (document 481 "CDSS Issues List 27062010"). The second more critical element is the logic defined within the core Decision Engine component. As a prelude to full clinical testing, iQudos developed their

own internal Test Platform to establish a mechanism to ensure that their development processes were test-driven.

The Decision Engine Test Platform makes it possible to quickly analyse a whole batch of patient data through the 'Decision Engine' component visit by visit and examine the outcomes. The Test Platform supports importing data in the ubiquitous CSV file format and so can include data provided by medical experts captured in templated Excel spreadsheets. The domain experts create 'ALL' possible disease profile permutations in order to exercise all the defined rules. They then seek to verify that the sample of synthetic patients and their symptoms is a sufficient sample to ensure that all the possible permutations are accounted for. The synthetic data are imported, analysed and results recorded by batch. The Test Platform quickly processes many patients' results without endlessly keying data into forms and allows the outcomes to be recorded and compared against predetermined and medically proven management. This becomes an iterative process of analyse data, examine and compare results, make rule adjustments and then repeat. The results are documented in document 498 "CDSS Decision Engine Test Platform Development Output Results (Stable Prostate Domain Module)".

The outcomes can be reviewed for any discrepancies by the domain experts, fine tuned and the whole test cycle repeated to ensure that one rule change does not affect any other paths. The outputs of these tests are captured in document 499 "CDSS Decision Engine Test Platform Exception Report" and document 500 "CDSS Decision Engine Test Platform Pre-Production Test Results".

c. Previous clinical experience

This is the first clinical investigation on the CDSS system for stable prostate cancer so there is no previous clinical experience with this device. However, previous clinical experience has been gained on a device with similar features, namely the LUTS CDSS system. A number of clinical studies, including in a pilot service, have been performed on this device (versions 0.1, 0.2 and 2.2), which like the device that is the subject of this clinical plan uses a software algorithm based on the applicable guidelines and is browser-based. These previous clinical investigations aimed to confirm that the device was performing correctly by providing advice that was clinically accurate and complied with the clinical management guidelines provided by BAUS, EAU and NICE. As the device consists of computer software, the priority was to make sure that the rules engine of the algorithm was functioning correctly, leading to the correct advice on management being generated in the advice letter to the GP. Initially synthetic data were run through the LUTS CDSS system, followed by prospective and retrospective studies, including a pilot study. In all of the clinical investigations, the outputs of the software (the advice letters) were reviewed by clinical experts in the field (urology specialists) for clinical accuracy and compliance with the guidelines. Errors were found

after both the synthetic patients (n=53) and first set of retrospective data (n=81). After correction of these, a larger retrospective study (n= 385, including the previous 134 patients) confirmed that for patients with varying presentations, the software was able to suggest appropriate management as no further errors were encountered in this study. Statistical methods were inappropriate as the management was correct or wrong and all patients had to be managed correctly. These studies were performed on LUTS CDSS version 0.1 and further detail can be found in documents 385, 386 and 387 in the Technical File.

Further clinical data were gained from two supervised pilot studies (Nottingham and Derby) of over two years duration. The LUTS CDSS system (v0.2) was used in a real clinical environment in GP practices by trained registered nurses and the company medical director, who is a trained consultant urological surgeon, provided clinical oversight. In the course of this study, 124 patients were seen, as both new and follow up patients. The data gave a good understanding of the use of the device in a realistic clinical environment and since there were no issues with either the software or its outputs, showed consistent performance with 100% accuracy during the course of the pilot studies (document 483 "Clinical Investigation Report: A Pilot Study of LUTS Clinical Decision Support Software (LUTS CDSS) System Version 0.2").

An additional clinical study was performed on the LUTS CDSS (v2.2) after updating following the issue of guidelines on LUTS from NICE. Some minor errors were found in the initial set of 124 patients run through this device and were corrected. A second set of 50 patients was then run and gave no further clinical errors, with agreement between the clinical experts and the device outcomes (document 484 "Clinical Investigation Report: A Clinical Evaluation of LUTS Clinical Decision Support Software (LUTS CDSS) System Version 2.2").

During the course of the clinical investigations on the similar LUTS CDSS device, there were no device-related adverse effects. Modifications were made during the earlier clinical studies as part of the development process of the software with no adverse effects on patients.

d. Device risk analysis and risk assessment

A risk analysis and assessment has been performed on the SPC CDSS version 3.0 device (document 495 "Stable prostate cancer clinical decision support software risk assessment") and the resulting risks are as follows:

1) Inappropriate patient management
2) Inability to access software
3) Unauthorised access to patient data

Of these risks, risk 1 is the clinically relevant risk of the device. This could be incorrect advice on the patient's management due to a logic error in the software, where the worst-case scenario would be a false negative such that a patient who should be referred to secondary care is missed. Despite the device's ability to follow the guidelines rigidly and to understand firmly its scope of practice, any single error has the possibility of being propagated to the care of multiple patients. To mitigate this, synthetic data have been run through the software to test the device for errors of logic and programming and ensure that patients can be managed error-free. These data have aimed to cover all likely scenarios of clinical presentation of SPC, and the clinical investigation described in this plan using actual clinical data will mitigate the risk for SPC CDSS version 3.0. Testing the software in all expected scenarios will minimise the risk of any unexpected outcomes and ensure that the correct clinically accurate outcomes are produced in line with the applicable guidelines. To further mitigate this risk, the outcomes of the first 500 patients will be reviewed as part of the post-market surveillance and a monthly review of 10% of the outcomes will be performed thereafter up to 2000 patients. Also as the users are trained nurses, unusual situations should be identified and can be verified with a urological expert. This risk can therefore be regarded as minor. If any changes are made to the applicable guidelines, the device will be modified accordingly and further testing performed to ensure compliance with the revised guidelines and continued clinical accuracy.

Another risk that could result in inappropriate patient management would be user interface error where incorrect information is input into the software. Although human error can never be ruled out, the risk is controlled both by having trained registered nurses as the CDSS users and also by the software design. Values input are validated to ensure that they are within acceptable normal ranges. For example, an IPSS score must be between 0 and 35. This risk is therefore minor.

Procedures have been put in place to compensate if the software is inaccessible. However this would only result in a short delay in providing the advice letter to the GP and is unlikely to be clinically significant. Similarly, procedures are in place to ensure that no unauthorised access to patient data occurs. Again, this risk is unlikely to result in any issues of clinical significance and both risks are minor.

As the risks associated with the device are all minor and expected to be rare occurrences, the balance of anticipated clinical benefit against the associated risks is positive, with the patient gaining the benefits of being treated in a community setting with a shorter waiting time for the consultation. The device also offers the benefit of recommending the best practice care as defined by the clinical guidelines as the device is programmed to always follow the guidelines.

As there is no medicinal product or invasive procedure involved, there is no risk of any possible interactions with concurrent medical interventions. No adverse device effects have been seen in patients to date and all modifications so far have had no effect on patients as they have been performed as part of device development and clinical investigations with urological consultant supervision.

3. Objectives of the Clinical Investigation

Hypothesis
The hypothesis is that the SPC CDSS v3.0 software can successfully manage male patients with stable prostate cancer in the community rather than a hospital setting and that the suggested management will be clinically accurate and compliant with the guidelines from specialty bodies NICE, BAUS and EAU.

Primary Objective
To test if the SPC CDSS version 3.0 device will produce clinical advice letters to the referring physician that will manage male patients with SPC safely and correctly in terms of recommendations on further investigation, referral and follow up in line with the guidelines from NICE, BAUS and EAU. This will verify the intended indication of the device of advising GPs on the management of stable prostate cancer.

Secondary Objective
To study the short and long term effects of the SPC CDSS v3.0 device by including both first visit and follow-up patient data. However, no follow up will be made on patients referred back to the GP or to secondary care.

A clinical investigation is necessary to ensure that the device performs as intended by providing management recommendations in letters to GPs that are clinically accurate and compliant with NICE, BAUS and EAU guidelines. This is the claim to be verified in the clinical investigation described in this plan. As the device consists of computer software, the priority is to make sure that the rules engine of the algorithm based on the clinical guidelines is functioning correctly, leading to the correct advice on management being generated in the advice letter to the GP. The main clinically relevant risk is that a logic error in the software would lead to incorrect management. This investigation will assess and reduce the chance of this risk by running retrospective data covering all of the presentations of SPC through the device (v3.0) until there is agreement between the management outcomes from the device and the management described in the clinical guidelines and considered reasonable by clinical experts (the investigators).

The hypothesis will be accepted when all of the retrospective clinical data sets can be run through the software without error, such that 100% of the patients are managed safely and correctly and in compliance with the applicable guidelines.

4. Design of the Clinical Investigation

The investigation described in this plan is a retrospective analysis of an anonymised database of data for patients with a diagnosis of stable prostate cancer. This investigation will use retrospective data run through the device to ensure that the rules engine of the algorithm based on the applicable clinical guidelines is functioning correctly, leading to the correct advice on management being generated in the advice letter to the GP. It is not necessary to recruit patients as the device is responding to data input by the user and retrospective data will be as accurate as prospective data in producing outcomes, without any risk to patients. Furthermore, the data have been captured in a clinical setting similar to the proposed normal conditions of use of the device. There is no control group as such as the device outcomes are being compared to the management outlined in the applicable guidelines rather than to another group of patients. However, the patient data will also be reviewed by clinical experts in SPC (the investigators) who will suggest appropriate outcomes in line with the applicable guidelines, and these will be compared to the device outcomes.

In order to minimize or avoid bias, each investigator (n=3) will each independently review the data for clinical accuracy and compliance with the applicable guidelines before seeing the outcome from the device. Any differences between the device and expert outcomes will be discussed as a panel to determine if amendments are required to the device. An independent expert (urologist) will also be included in the panel discussions.

Primary Endpoint

The agreement between clinical experts in SPC and the device on the safe and appropriate management of SPC described in the advice letters produced by the device, taking into account adherence with the applicable guidelines. Agreement must be reached on all letters (100% agreement) to achieve the primary endpoint and agreement must be reached on the last subset of letters after no further modifications have been made to the device.

Secondary Study Endpoints

1) The assessment of long term effects of the device in managing patients with data from multiple visits. Advice letters will be generated for the latest visit presented and will take into account any previous visit data in the assessment and recommended management of the patient. The advice should be appropriate, clinically accurate and comply with the applicable guidelines taking into account all data from the previous

visits. This will provide evidence that the device is able to manage patients successfully over time.

2) The matching of the software decision letters for patients for whom the same management action is required. All letters with the same management outcome should match. This will ensure consistency of the outcomes between patients and visits.

a. Methods

The variable to be measured in this investigation is the outcome from the SPC CDSS system of an advice letter to the patient's GP advising on the appropriate clinical management of the patient. The review of these letters independently by clinical experts (the investigators) and non-clinical reviewers where appropriate will allow both the primary and secondary endpoints to be met. There is no test equipment required for assessment of the study variables.

The investigation aims to test that the device (SPC CDSS v3.0) produces results that are consistent with the management outcomes from the applicable clinical guidelines and clinically accurate. The method for the investigation is described below.

1) A DHFT IT expert, who will be independent of the research team, will extract the retrospective patients' data from the databases of clinical data held in the urology department of Royal Derby Hospital.
2) The data will be anonymised by the IT expert, who has access to patients' data including their identifiable details, by using a unique Study ID for each patient and the anonymised data will be provided to the investigators. A link between the study ID and the patients' details will be maintained by the IT expert, and not provided to the investigators, for traceability and for historic record detail comparison.
3) The data will be reviewed to ensure that all likely presentations of SPC are present and to exclude any data that do not meet the inclusion criteria. If additional data are required, then a research assistant, who will be independent of the research team and has access to patients' clinical identifiable details on a regular basis, will be required to capture this data from clinical sources such as patient health records and these data will be anonymised by the IT expert and added to the existing electronic database. A link between the study ID and the patients' details will be maintained by the IT expert, and not provided to the investigators, for traceability and for historic record detail comparison.
4) A subset of retrospective anonymised data from all likely presentations of SPC (n=100 patients) from the urology clinical database will be run through the SPC CDSS v3.0 device.
5) For each patient (n=100), the data will be printed as a letter without the recommendation for treatment and supplied to the investigators for review.

6) These letters will then be independently reviewed by each of the investigators (n=3). The advice letter produced as the outcome contains all the relevant data required by the investigator to reach a conclusion on the most appropriate clinical management of the patient in compliance with the applicable guidelines, and includes data from previous visits if appropriate.
7) The investigator will classify the patient according to their history, presenting conditions and clinical management recommendation, taking into account adherence with the applicable guidelines, by recording the appropriate letter codes on the Case Report Form (CRF) against the Subject ID number. The list of possible letter codes is provided as part of the CRF and also illustrated in Appendix Two. All relevant letter codes should be included for each patient. Comments can also be added for each patient in the "Comments" section of the CRF.
8) The Clinical Monitor will tabulate the outcomes from the device and from the investigators into an Excel spreadsheet to allow comparison of the device outcomes with the investigators' outcomes.
9) The device recommendations will then be compared to the recommendations from the investigators and the percentage of recommendations that agree for each investigator and overall will be noted.
10) The device outcomes will then be shared with the investigators and the outcomes discussed as a panel, which will include an additional independent expert.
11) If the device recommendation complies with the applicable guidelines and is clinically accurate (all the investigators and the device agree), then the case will be deemed a Pass and the advice letter for the next patient will be compared until every patient case has been reviewed and passed.
12) If all patients are assessed as a Pass (n=100 patients) after the first run through the device, then further patients will be analysed as described in step 17 below.
13) If a discrepancy in the outcome between the investigators and the device occurs, then the panel will discuss the outcome and the most appropriate and correct management agreed, taking into account adherence with the applicable guidelines. The decision of the panel must be unanimous and discussions must continue until this is achieved.
14) If the device was at fault then the case will be deemed a Fail and this will be recorded on the DCF. If the investigators conclude that the device was correct, it will be deemed a Pass.
15) The software will then be corrected and the entire patient subset (n=100) will be retested through the device from the start of the subset.
16) A Clinical Monitor will then audit the device outcomes against the outcomes agreed by the panel of investigators. If 100% agreement has been achieved then the subset has passed; if not, then further corrections will be made to the software, the patient subset retested and the outcomes audited until 100% agreement is achieved.

17) A further subset of 100 patients will then be run through the device with the process being repeated from step 4.
18) When the device recommended outcomes for all patients in all of the subsets tested Pass so that there is 100% agreement between the investigators and the device, then this will achieve the primary endpoint. If 100% agreement is achieved from the first run of the second subset of 100 patients, the true error will be estimated without analysing more patients (n=52).
19) If there are any Fails after the second run through of the 100 patients, then modifications will be made and both subsets of patients will be retested until correct (n=200). A third random subset from the remaining retrospective anonymised data (n=52 patients) from the urology clinical database will then be analysed by the SPC CDSS v3.0 device repeating the procedure from step 4 onwards for estimating the true error of the software. If any Fails are found, modifications will be made in the software and all patients retested until they Pass. Further subsets will then be tested until all previously tested patients Pass and the latest subset of 52 patients passes on the first analysis by the device. This will meet the primary endpoint.
20) The above process will also ensure that where possible, the performance of the device over a longer term has been assessed by analysing patients with more than one visit. If all patient outcomes have passed then secondary endpoint 1 assessing longer term has also been met.
21) A random subset of 50 patients from the first 100 patients analysed will be run through the device again through the user interface by typing in the data manually to simulate normal conditions of use.
22) The outcomes from the device from step 21 will then be compared with the device outcomes from the automated run through the device (after any corrections from the clinical expert outcomes) to confirm that the user interface is correctly communicating with the rules engine of the device. If any discrepancies occur, the device will be corrected and the remaining 50 patients from the first 100 will be run and this step repeated.
23) A review will be performed by a non-clinical reviewer to classify the advice letters (n=200) into groups having the same outcome. These groups of letters will be used to assess if the device produces identical letters for patients with the same outcome classification.
24) If there are discrepancies in the letter text, then the software will be checked and the process repeated until the letter content is clinically acceptable. This will meet secondary endpoint 2.

Inclusion Criteria

The records of anonymised male patients diagnosed with stable prostate cancer that include a full record of all of the data required by the software.

Exclusion Criteria
Any anonymised patient's record without the full data set as defined in the inclusion criteria.

As anonymised data are being used, there is no point of enrolment into the investigation and no subjects to undergo any procedures. Thus there is no need for any criteria or procedures for withdrawal and discontinuation of subjects. Similarly as anonymised patient data are being used, there are no adverse events or adverse device effects as there are no subjects being treated. The only foreseeable factor that may compromise the outcomes from the device or interpretation of results is if patient cases arise that are not expressly covered in the guidelines.

The retrospective data in this study are taken from a study population of 325 patients. Initially the sample size of patients in this study is 200 patients, analysed in two subsets of 100 patients. It is believed that this sample size is adequate to test the software, as it will cover all expected clinical presentations of SPC. If 100% agreement is not reached between the investigators and the device after analysis of the second subset (n=100) by the SPC CDSS v3.0 system at the first run, then a further random subset of 52 patients will be analysed for estimating the true software error.

As the device is used from the secure N3 server with a single point of access, there is one device used in this investigation. After the clinical investigation has been successfully completed, this device will continue to be used in clinical practice until a new version becomes necessary.

5. Statistical Considerations

The hypothesis is that the SPC CDSS v3.0 device can successfully manage male patients with stable prostate cancer in the community rather than in a hospital setting and that the suggested management will be clinically accurate and comply with the guidelines from specialty bodies NICE, BAUS and EAU. If the investigation described in this clinical investigation plan is successful, then this hypothesis will be accepted and the device will fulfil its intended indication.

The statistical analysis will be undertaken by the DHF Trust statistician using the Stata v11.0. A sample of 200 cases of anonymised data representing all possible management categories of SPC, in two subsets of 100 patients, will be used for the validation/correction of the software against the true clinical decision for achieving error rate of 0%. If the first subset of 100 patients fails then the precision will be ±1.1% for the 95% Confidence Interval of error rate of 0.3%, with power 80%. If errors are not found in the second subset of 100 patients during their first run, then the precision of 0% true error rate will be ±3%

for the 95% Confidence Interval, with power 80%. If errors are found in the second subset of 100 patients, then the device will be corrected for each error and rerun in the 200 patients until error-free. This would require a random sample of a further 52 different patients to be run through the device for estimating the true error rate. This sample of patients can produce, with 80% power, a 95% Confidence Interval of 0.3% error rate with precision +/-1.5%. The statistics will not change further. The power at 80% is reasonable for a clinical investigation of this type. If the error rate of 0.3% is observed, this can be minimised in future with post-marketing surveillance data and the true error rate will be re-estimated.

The outcome from the investigators will be compared to that from the device for estimating their inter rater agreement using the Kappa coefficient (95% Confidence Interval). The classification of the patients from three different clinicians based on the device letter will be compared using the Kappa coefficient (95% Confidence Interval). The percentage of agreement between the investigators and the device will be recorded after each subset of patients. To be correct, the management of all patients must be clinically accurate and reasonable and comply with the applicable guidelines as judged by three investigators. This would constitute a Pass. If there is not agreement in the management of all patients after discussion of the investigators as a panel, then the device is not managing patients correctly and this would constitute a Fail. The management should be correct in 100% of cases. The sensitivity, specificity, positive and negative predicted values (95% Confidence Interval) will be estimated for the accuracy of the software decision against the correct clinical decision.

The number of times that the software is amended to achieve 100% correct management will be recorded. Any missing data will be noted on the CRF along with the patient ID numbers. Unused data will be recorded in case further testing is required. Any deviation(s) from this statistical plan shall be described and justified in the final report.

6. Deviations from the Clinical Investigation Plan

Any deviation from the CIP shall be recorded together with an explanation for the deviation. Deviations will be reported to the sponsor and the significance of the deviation assessed.

As there are no subjects, there are no issues with withdrawal and discontinuation.

7. Amendments to the Clinical Investigation Plan

All amendments to the CIP shall be agreed between the sponsor and the clinical investigator(s) and be recorded with a justification for the amendments. Deviations will be reviewed to determine the need to amend the CIP or terminate the investigation.

A list of the clinical investigators will be maintained by the sponsor and available throughout the trial. The definitive list of all the investigators will be provided with the final report.

8. Adverse Events and Adverse Device Effects

Not applicable as subjects will not be recruited.

9. Early Termination or Suspension of the Investigation

Not applicable as subjects will not be recruited.

10. Publication Policy

Publication will be sought in relevant Professional Peer Reviewed Journals and it is hoped that the outcome can be presented at specialty clinical meetings to be decided.

11. Case Report Forms (CRF)

The CRF is attached in Appendix One and the DCF in Appendix Three. These forms will be the primary data collection instruments for recording the results of this clinical investigation. All data requested on the CRF and DCF will be recorded. All missing data will be explained. All entries will be printed legibly in indelible black ink. If any entry error has been made, a single straight line will be drawn through the incorrect entry and the correct data will be entered above it. All such changes will be initialled and dated. Errors will not be erased by any method. For clarification of illegible or uncertain entries, the clarification will be printed above the item, then initialled and dated.

12. Monitoring Arrangements

The investigators will permit study-related monitoring, audits and inspections by the Sponsor, regulatory inspectors and the Clinical Monitors. This study will be monitored in accordance with the relevant parts of BS EN ISO 14155-1:2009 and DHFT procedure SOP-

RGE-08. In line with the Research Governance Framework, the investigators and software developers will ensure that the Clinical Monitors or other regulatory monitoring authority are given access to all source data and investigation-related documents and investigation-related facilities. The anonymised data used in the investigation will be verified against the original source data. The data presented to the investigators in the advice letters will be audited against the anonymised data for completeness and consistency with the anonymised data.

13. Data and Quality Management

Database Management
Information about investigation subjects in the original database will be kept confidential and managed according to the requirements of the Data Protection Act, NHS Caldicott Guardian, the Research Governance Framework for Health and Social Care, Ethics Committee approval, IgSoc documents, DHFT IM&T policy and iQudos Company Policy as defined in document 379 "Quality Management System Manual". The Company Medical Director, Michael Henley, will be the "Custodian" of the data and subjects will not be identifiable from any future publications.

The applications data currently resides on a high specification server physically located in a secure data centre at the Royal Derby Hospital and is only available to the NHS and authorised commercial third parties. Operating within the N3 Connecting For Health environment immediately implies a minimum level 2 conformance of the Information Governance Statement of Compliance. This is the NHS framework that provides the highest standards of Information Governance relating to patient records. It is the standard that MS3 Ltd (aka iQudos) and DHFT adhere to. The software developers were contracted to a company that has level 2 IGSoC (Information Governance Statement of Compliance) during the development of the software and are now working for the MS3 Ltd (aka iQudos).

Treatment of Data
The source data will be the existing electronic databases containing details of patients' clinical progress and the patients' health records. All patients' records will be completely anonymised using a study ID, as patient's identifiable information is not needed for this investigation. Removing any data element that could uniquely identify the patient, including but not limited to their NHS number, their name, and their address and telephone number, will anonymise the data. The database primary identity key is maintained by the IT expert, who is independent of the investigators, to identify the patient for the purposes of the investigation for traceability and for historic record detail comparison. The identity key will not be available to the investigators. The anonymised data are used on internal test

platforms and these are the repositories used to present the data to the CDSS algorithms. The clinical data are iterated through the SPC CDSS v3.0 device as described above in section 2.2 to produce clinical guidance in the form of a GP letter that is made available to the investigators.

All paper clinical documentation, including CRFs and DCFs, will be kept in a secure locked filing cabinet at MS3 Ltd (aka iQudos) office. The data from the CRFs will be input into an Excel spreadsheet for comparison with the device outcomes.

Access to the anonymised data will be for the research team consisting of investigators/ clinical experts, clinical project manager, clinical monitors, non-clinical reviewers and the software developers.

Source Data Verification

Once the anonymised data has been transferred onto the N3 server, it will be verified for consistency and completeness with the source data. The Clinical Monitor will do this by manually auditing the data on the N3 server against the source data.

After the advice letters have been produced, the data in these letters will be manually audited against the anonymised data on the N3 server for each patient record as indicated by the unique study ID.

Data archiving

Data will be archived by Steven Pike (Risk Officer: IGSoC) at MS3 Ltd (aka iQudos).

Retention Period

All the investigation documents and the anonymised databases will be retained for at least ten years beyond the conclusion of the investigation in line with DHFT governance guidelines.

Indemnity

NHS indemnity will apply in the event of a claim by, or on behalf of, participants for negligent harm. There will be no special arrangements for non-negligent harm but the normal NHS Complaints mechanism will be available to all participants.

Funding Source

The study is funded by Derby Hospitals NHS Foundation Trust as part of matched funding for the EMDA Demonstration Grant EMX0593 "Clinical Decision Support System".

14. References

1) NICE guideline: CG97 Lower Urinary Tract Symptoms, May 2010, http://guidance.nice.org.uk/CG97.

2) Speakman MJ. *Guideline for the primary care management of male lower urinary tract symptoms.* BJU Int (2004) 93:985-990.

3) Madersbacher S., Alivizatos G., Nordling J., Sanz C.R., Emberton M., de la Rosette J.J. *EAU 2004 guidelines on assessment, therapy and follow-up of men with lower urinary tract symptoms suggestive of benign prostatic obstruction (BPH guidelines).* Eur Urol 46 (2004) 547 - 554.

4) John R., Sanders P., Chilton C. P., Henley M. J. *A prospective study on therapeutic advice provided by an expert computing system.* In conference proceedings of The British Prostate Group Autumn Meeting (2006).

5) John R., Jackson B., Sanders P., Chilton C.P., Henley M.J. *Can a computer kick the urologist out of the office?* BJU Int 99 (2007) 42.

6) Rajkumar G.N., Small D.R., Conn I.G. *Computerised triage in a prostate assessment clinic.* Prostate Cancer and Prostatic Diseases 7 (2004) 118-121.

7) Norg R.J.C., Portegijs P.J.M., van de Beek K., van Schayck O. Knottnerus J.A. *A decision aid for GPs for the treatment of elderly male patients with lower urinary tract symptoms (LUTS).* Family Practice 23 (2006) 666-673.

8) Rios M., Desandes E., Bresson B., Klein I., Lesur A., Boisson F., Demange V., Bey P. *Clinical practice guidelines in cancerology: comparative study of three decision support-systems for breast and prostate cancer in Lorraine French region.* Bull Cancer 90 (2003) 363-70.

9) Daniyal A., Abidi S.R., Abidi S.S. *Computerizing Clinical Pathways: Ontology-Based Modeling and Execution.* Studies in Health Technology & Informatics 150 (2009) 643-7.

10) Abbod M.F., Catto J.W.F, Linkens D.A., Hamdy F.C. *Application of Artificial Intelligence to the Management of Urological Cancer.* J. Urol. 178 (2007) 1150-1156.

15. Appendices

This section contains all pertinent documents associated with the management of the study.

Appendix One – Case Report Form (CRF)
Appendix Two – Mind map showing classification codes for possible outcomes
Appendix Three – Data Collection Form (DCF)
Appendix Four – Literature Review Protocol for Stable Prostate Cancer Clinical Decision Support Software System
Appendix Five – Literature Review Report for Stable Prostate Cancer Clinical Decision Support Software System

Appendix One

SPC CDSS Clinical Investigation Case Report Form (CRF)

Date of Investigation:

Place of Investigation: DHFT

Investigation Identification No.: RD-5103-006-09

CIP Document and Version No.: 497, version 1.0

Device: SPC CDSS version 3.0

Medical Diagnosis: Stable prostate cancer (SPC)

Inclusion/Exclusion Criteria: All data have been reviewed for compliance with the inclusion/exclusion criteria. All data presented here comply.

Procedural Data

1) The procedure for review of the device outcomes will be followed as outlined in document number 497, version 1.0, section 4.1 Methods.
2) Record the appropriate letter codes from Appendix Two for each subject in the box next to the Subject Study ID No. for that subject. Double check that the correct codes have been written next to the correct ID No.
3) Any additional information for each subject can be recorded in the Comments box next to the subject referred to.
4) Please record codes for all Study ID Numbers.
5) Please sign and date the document in the box provided.

Since the live data extract from Derby Hospitals is focused primarily on just PSA monitoring, synthetic abnormalities have been added to selected patients in order to validate all possible conditions and only when clinically necessary data are missing. The purpose of this is to fully test the device in cases where data have not been recorded clinically.

Subject Study ID No.	SPC Management Codes	Comments
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		

Subject Study ID No.	SPC Management Codes	Comments
19		
20		
21		
22		
23		
24		
25		
26		
27		
28		
29		
30		
31		
32		
33		
34		
35		
36		
37		

Subject Study ID No.	SPC Management Codes	Comments
38		
39		
40		
41		
42		
43		
44		
45		
46		
47		
48		
49		
50		
51		
52		
53		
54		
55		
56		

Subject Study ID No.	SPC Management Codes	Comments
57		
58		
59		
60		
61		
62		
63		
64		
65		
66		
67		
68		
69		
70		
71		
72		
73		
74		
75		

Subject Study ID No.	SPC Management Codes	Comments
76		
77		
78		
79		
80		
81		
82		
83		
84		
85		
86		
87		
88		
89		
90		
91		
92		
93		
94		

Subject Study ID No.	SPC Management Codes	Comments
95		
96		
97		
98		
99		
100		

CLINICAL INVESTIGATOR	
SIGN:	DATE ☐☐ / ☐☐☐ / ☐☐
PRINT NAME:	

Subject Study ID No.	SPC Management Codes	Comments
101		
102		
103		
104		
105		
106		
107		
108		
109		
110		
111		
112		
113		
114		
115		
116		
117		
118		

Subject Study ID No.	SPC Management Codes	Comments
119		
120		
121		
122		
123		
124		
125		
126		
127		
128		
129		
130		
131		
132		
133		
134		
135		
136		
137		

Subject Study ID No.	SPC Management Codes	Comments
138		
139		
140		
141		
142		
143		
144		
145		
146		
147		
148		
149		
150		
151		
152		
153		
154		
155		
156		

Subject Study ID No.	SPC Management Codes	Comments
157		
158		
159		
160		
161		
162		
163		
164		
165		
166		
167		
168		
169		
170		
171		
172		
173		
174		
175		

Subject Study ID No.	SPC Management Codes	Comments
176		
177		
178		
179		
180		
181		
182		
183		
184		
185		
186		
187		
188		
189		
190		
191		
192		
193		
194		

Subject Study ID No.	SPC Management Codes	Comments
195		
196		
197		
198		
199		
200		

CLINICAL INVESTIGATOR	
SIGN:	DATE ☐☐ / ☐☐☐ / ☐☐
PRINT NAME:	

Subject Study ID No.	SPC Management Codes	Comments
201		
202		
203		
204		
205		
206		
207		
208		
209		
210		
211		
212		
213		
214		
215		
216		
217		
218		
219		

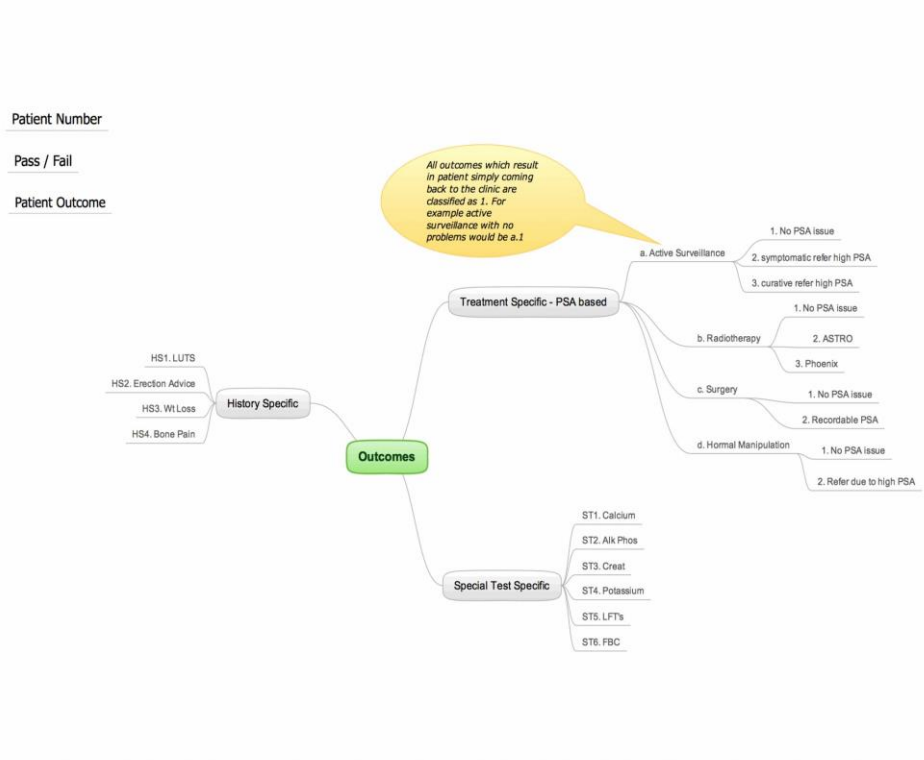
Subject Study ID No.	SPC Management Codes	Comments
220		
221		
222		
223		
224		
225		
226		
227		
228		
229		
230		
231		
232		
233		
234		
235		
236		
237		
238		

Subject Study ID No.	SPC Management Codes	Comments
239		
240		
241		
242		
243		
244		
245		
246		
247		
248		
249		
250		
251		
252		

CLINICAL INVESTIGATOR	
SIGN:	DATE ☐☐ / ☐☐☐ / ☐☐
PRINT NAME:	

Appendix Two

Mind map showing classification codes for possible outcomes



Appendix Three

SPC CDSS Clinical Investigation Data Collection Form (DCF)

Date of Investigation Meeting:

Place of Investigation: DHFT

Investigation Identification No.: RD-5103-006-09

CIP Document and Version No.: 497, version 1.0

Device: SPC CDSS version 3.0

Medical Diagnosis: Stable prostate cancer (SPC)

Inclusion/Exclusion Criteria: All data have been reviewed for compliance with the inclusion/exclusion criteria. All data presented here comply.

Procedural Data

1) The procedure for review and discussion of the device outcomes will be followed as outlined in document number 497, version 1.0, section 4.1 Methods.
2) Record the result (Pass or Fail) for each subject in the box next to the appropriate Subject Study ID No. for that subject. Double check that the correct result has been written next to the correct ID No.
3) Any additional information for each subject can be recorded in the Comments box next to the subject referred to.
4) Please record codes for all Study ID Numbers.
5) Please sign and date the document in the box provided.

Subject Study ID No.	CDSS management recommendation PASS or FAIL	Comments
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		

Subject Study ID No.	CDSS management recommendation PASS or FAIL	Comments
20		
21		
22		
23		
24		
25		
26		
27		
28		
29		
30		
31		
32		
33		
34		
35		
36		
37		
38		

Subject Study ID No.	CDSS management recommendation PASS or FAIL	Comments
39		
40		
41		
42		
43		
44		
45		
46		
47		
48		
49		
50		
51		
52		
53		
54		
55		
56		
57		

Subject Study ID No.	CDSS management recommendation PASS or FAIL	Comments
58		
59		
60		
61		
62		
63		
64		
65		
66		
67		
68		
69		
70		
71		
72		
73		
74		
75		
76		

Subject Study ID No.	CDSS management recommendation PASS or FAIL	Comments
77		
78		
79		
80		
81		
82		
83		
84		
85		
86		
87		
88		
89		
90		
91		
92		
93		
94		
95		

Subject Study ID No.	CDSS management recommendation PASS or FAIL	Comments
96		
97		
98		
99		
100		

CLINICAL INVESTIGATOR	
SIGN:	DATE ☐☐ / ☐☐☐ / ☐☐
PRINT NAME:	

Subject Study ID No.	CDSS management recommendation PASS or FAIL	Comments
101		
102		
103		
104		
105		
106		
107		
108		
109		
110		
111		
112		
113		
114		
115		
116		
117		
118		

Subject Study ID No.	CDSS management recommendation PASS or FAIL	Comments
119		
120		
121		
122		
123		
124		
125		
126		
127		
128		
129		
130		
131		
132		
133		
134		
135		
136		
137		

Subject Study ID No.	CDSS management recommendation PASS or FAIL	Comments
138		
139		
140		
141		
142		
143		
144		
145		
146		
147		
148		
149		
150		
151		
152		
153		
154		
155		
156		

Subject Study ID No.	CDSS management recommendation PASS or FAIL	Comments
157		
158		
159		
160		
161		
162		
163		
164		
165		
166		
167		
168		
169		
170		
171		
172		
173		
174		
175		

Subject Study ID No.	CDSS management recommendation PASS or FAIL	Comments
176		
177		
178		
179		
180		
181		
182		
183		
184		
185		
186		
187		
188		
189		
190		
191		
192		
193		
194		

Subject Study ID No.	CDSS management recommendation PASS or FAIL	Comments
195		
196		
197		
198		
199		
200		

CLINICAL INVESTIGATOR	
SIGN:	DATE ☐☐ / ☐☐☐ / ☐☐
PRINT NAME:	

Subject Study ID No.	CDSS management recommendation PASS or FAIL	Comments
201		
202		
203		
204		
205		
206		
207		
208		
209		
210		
211		
212		
213		
214		
215		
216		
217		
218		
219		

Subject Study ID No.	CDSS management recommendation PASS or FAIL	Comments
220		
221		
222		
223		
224		
225		
226		
227		
228		
229		
230		
231		
232		
233		
234		
235		
236		
237		
238		

Subject Study ID No.	CDSS management recommendation PASS or FAIL	Comments
239		
240		
241		
242		
243		
244		
245		
246		
247		
248		
249		
250		
251		
252		

CLINICAL INVESTIGATOR	
SIGN:	DATE ☐☐ / ☐☐☐ / ☐☐
PRINT NAME:	

APPENDIX SIX – AMENDMENT TO CLINICAL INVESTIGATION PLAN

Introduction

On gaining access to the urology database at Derby Hospital containing the data for the investigation, it was found not to be as comprehensive as required to fully test the device. Although there are sufficient patients, there is not the breadth of scenarios. As detailed in the clinical investigation plan, more data were sought from the records. However in some cases the required data are not accessible, not done or not recorded and there are certain blood results and symptoms missing. The inclusion criteria states that records must include "a full record of all of the data required by the software". Unfortunately this would leave few patients to test. The inclusion criteria have therefore been revised as follows:

Inclusion Criteria

The records of anonymised male patients diagnosed with stable prostate cancer that include more than one follow-up appointment with an associated PSA reading taken less than or equal to 60 days before the appointment date.

Since the live data extract from Derby Hospitals is focused primarily on just PSA monitoring, synthetic abnormalities have been added to selected patients in order to validate all possible conditions and only when clinically necessary data are missing. The purpose of this is to fully test the device in cases where data have not been recorded clinically. Normal readings will also be added to the remaining patients so as not to introduce bias into the investigation.

ON BEHALF OF THE SPONSOR, TERESA GRIEVE
SIGN DATE ☐☐ / ☐☐☐ / ☐☐
TERESA GRIEVE, ASSISTANT DIRECTOR OF R&D, DERBY HOSPITALS NHS FOUNDATION TRUST, ROYAL DERBY HOSPITAL

AUTHOR AND CLINICAL INVESTIGATION MONITOR/PROJECT MANAGER
SIGN DATE ☐☐ / ☐☐☐ / ☐☐
SUE HAGAN, HEAD OF CLINICAL AND REGULATORY AFFAIRS, IQUDOS

CONFIDENTIAL: THIS DOCUMENT IS THE PROPERTY OF DERBY HOSPITALS NHS FOUNDATION TRUST, ROYAL DERBY HOSPITAL

PRINCIPAL CLINICAL INVESTIGATOR
SIGN DATE ☐☐ / ☐☐☐ / ☐☐
STEPHEN THOMAS, CONSULTANT UROLOGIST, DERBY HOSPITALS NHS FOUNDATION TRUST, ROYAL DERBY HOSPITAL

Appendix 2: Refined decision rules of the prototype

Rules for Stable Prostate Cancer Clinic

The rule set that the system uses is based on the Derby Nurse Led Clinic protocol. I thought that it would be a good idea to very briefly list out the rules for investigators as there is some translation from a hospital based system, for example discuss with consultant becomes refer to hospital.

Cancer Management Pathway

Radical Prostatectomy
Any measurable PSA = refer to hospital

Radiotherapy
Failure on ASTRO or Phoenix = refer to hospital

Hormonal Manipulation
PSA>20 refer to hospital *(or local value)*

Active surveillance
State that patient should be followed up at hospital as not suitable for community follow up.

Watchful Waiting
refer if PSA >20 (or locally set maxima)
If [∂PSA+current PSA > PSA max]. In other words previous rise in PSA added to current PSA greater than threshold (threshold to be set by location) then see at 3 months rather than 6.

Patient Specific PSA Threshold (new rule)
Consultants can specify directly a future PSA value that should trigger re-referral, for example patients who might benefit from chemotherapy. If PSA >threshold then states patient should be referred

History Specific Problems
LUTS -advice given to patient, noted to GP
ED- -advice given to GP for management
Wt Loss - note made, GP to consider investigating
Bone Pain - reported for action see below

Missing Blood Tests
Switchable between brought back or arrange tests and results to GP

Abnormal Blood Tests Other Than PSA

Should be recorded in text

Management of Blood abnormalities

- Ca2+ high -> refer to hospital
- Alk Phos high – Consider investigation in primary care (state that differential diagnosis can be secondaries).
- Ca2+ and Alk Phos high -> Referral to hospital
- Creatinine high -> consider investigating/referral to hospital
- K+ high -> isolated mention and repeat
- K+ low -> report only
- Creatinine and K+ high -> urgent referral to hospital
- ALT high on hormones -> urgent refer GP to consider investigation
- ALT high not on hormones - report to GP
- Hb – report to GP

Management of Blood and Clinical Abnormalities

- Bone pain alone -> Consider referral and investigation
- Bone pain + Ca2+ -> Refer to hospital
- Bone pain + alk phos -> Refer to hospital to exclude mets
- Bone pain + alk phos + hypercalcaemia -> refer to hospital

Post Radiotherapy

Patients need flexible sigmoidoscopy every 5 years following radiotherapy

Appendix 3: Service evaluation approval

RE: RE: Ethics approval

Shone Angela

Wed 06/05/2015 10:27

To: Salem Hesham <mzxhs@exmail.nottingham.ac.uk>;

Dear Hesham,

This is a service evaluation of an existing system and not clinical research that requires an ethical opinion. The data being supplied is anonymised so there are no ethical implications for the patients' data involved.
All you need therefore is the Trust's approval to access the database.

Best wishes
Angie

From: Salem Hesham
Sent: 04 May 2015 21:30
To: Shone Angela
Subject: RE: Ethics approval

Dear Angela,

Thank you for agreeing to review our protocol for validating the prostate cancer Expert System. Please let me know if ethics approval is mandatory.

Kind regards,

Hesham

Mobile: 07726615815

SALEM, Hesham (NORTHAMPTON GENERAL HOSPITAL NHS TRUST)

From: Stanford Richard (DERBY TEACHING HOSPITALS NHS FOUNDATION TRUST)
Sent: 16 June 2015 15:43
To: Salem Hesham (DERBY TEACHING HOSPITALS NHS FOUNDATION TRUST); Badhan Ranjit (DERBY TEACHING HOSPITALS NHS FOUNDATION TRUST)
Subject: RE:

Hi Hesham

Of course I am happy to approve this as Urology Departmental Audit Lead. Is there any particular paperwork that I need to sign for this given that it is classed as Service Evaluation?

Richard

Mr Richard Stanford
Consultant Urological Surgeon
Royal Derby Hospital

_____

From: Salem Hesham (DERBY TEACHING HOSPITALS NHS FOUNDATION TRUST)
Sent: 08 June 2015 17:07
To: Stanford Richard (DERBY TEACHING HOSPITALS NHS FOUNDATION TRUST)
Cc: Badhan Ranjit (DERBY TEACHING HOSPITALS NHS FOUNDATION TRUST)
Subject:

Dear Richard,

I have been in touch with our audit department and Ranjit (one of our Clinical Audit & Effectiveness Facilitators) confirmed that my study proposal falls in the category of service evaluation rather than a clinical. This would be keeping with the opinion that I have received from the University and the Hospital R&D. In that case it would only require your approval from yourself as the audit lead. It would include a similar study to the one that you have kindly contributed to but with aims to validate the software not just the knowledge. It will still be anonymised with no effect on the patients ' care or service delivery. I have attached a copy of the protocol and would be happy to discuss any further details.

Kind regards,

Hesham

RE: R&D approval

Patel Ramila (DERBY TEACHING HOSPITALS NHS FOUNDATION TRUST) <ramila.patel8@nhs.net>

Thu 28/05/2015 12:55

To:Salem Hesham <mzxhs@exmail.nottingham.ac.uk>;

Dear Hesham,
Many apologies again for not responding sooner.

As per our discussion, and your email below, your project is Service Evaluation. As such you will not need R&D approval from this office.

Good luck!

Kind regards
Ram

Dr Ramila Patel
Research Governance & Clinical Trials Manager

Research and Development Office
Royal Derby Hospital
Room 5033, University of Nottingham GEM School
Derby Teaching Hospitals NHS Foundation Trust
Uttoxeter Road
Derby, DE22 3DT

Tel: 01332 789339
E-mail: Ramila.patel8@nhs.net

All general enquiries for the R&D Department should be directed to our admin support address dhft.RandDadmin@nhs.net where we will endeavour to provide a same day response service



Did you know that you have 70 days to consent a research participant from the date of valid application submission? The R&D team can help you meet your study's targets. Contact us at dhft.randdadmin@nhs.net

'Reproduced with the kind permission of University Hospitals of Leicester NHS Trust'

From: Hesham Salem [mailto:mzxhs@nottingham.ac.uk]
Sent: 11 May 2015 16:33
To: Patel Ramila (DERBY TEACHING HOSPITALS NHS FOUNDATION TRUST)
Subject: FW: R&D approval

Dear Ramila,

I was given your contact from Trish to ask for advice with regards to applying for R&D approval. I am a PhD student in the University of Nottingham. One of my thesis studies includes validating an expert system software in an existing service running in the Urology outpatient. I will be only using anonymised data and I will not be interacting with any patient. The data will only include routine information and I will not be seeking any extra. The protocol has been reviewed by the

University Ethics (Angela Shone) and they have classified the study as service evaluation (please find below their response) and they suggested R&D approval.
I am based in the Hospital and would be glade to come over to your office and explain this further. I have also attached a copy of my protocol.

Kind regards,

Hesham

From: Salem Hesham
Sent: 06 May 2015 16:46
To: 'teresa.grieve@nhs.net'
Subject: R&D approval

Dear Teresa,

I hope you are keeping well. I am wondering if our R&D can approve a validation study of the hospital prostate cancer software. This would be part of my PhD and will only include analysing existing data which will be completely anonymised. The protocol has been reviewed by the University ethics and was classified as service evaluation and ethics approval was deemed unnecessary.

I have attached a copy of the protocol and would be happy to clarify further as required.

Kind regards,

Hesham

From: Shone Angela
Sent: 06 May 2015 10:28
To: Salem Hesham
Subject: RE: RE: Ethics approval

Dear Hesham,

This is a service evaluation of an existing system and not clinical research that requires an ethical opinion. The data being supplied is anonymised so there are no ethical implications for the patients' data involved.
All you need therefore is the Trust's approval to access the database.

Best wishes
Angie

From: Salem Hesham
Sent: 04 May 2015 21:30
To: Shone Angela
Subject: RE: Ethics approval

Dear Angela,

Thank you for agreeing to review our protocol for validating the prostate cancer Expert System. Please let me know if ethics approval is mandatory.

Kind regards,

Hesham

Mobile: 07726615815

This message and any attachment are intended solely for the addressee
and may contain confidential information. If you have received this
message in error, please send it back to me, and immediately delete it.

Please do not use, copy or disclose the information contained in this
message or in any attachment. Any views or opinions expressed by the
author of this email do not necessarily reflect the views of the
University of Nottingham.

This message has been checked for viruses but the contents of an
attachment may still contain software viruses which could damage your
computer system, you are advised to perform your own checks. Email
communications with the University of Nottingham may be monitored as
permitted by UK legislation.

****
****

This message may contain confidential information. If you are not the intended recipient
please inform the
sender that you have received the message in error before deleting it.
Please do not disclose, copy or distribute information in this e-mail or take any action in
reliance on its contents:
to do so is strictly prohibited and may be unlawful.

Thank you for your co-operation.

NHSmail is the secure email and directory service available for all NHS staff in England and
Scotland
NHSmail is approved for exchanging patient data and other sensitive information with NHSmail
and GSi recipients
NHSmail provides an email address for your career in the NHS and can be accessed anywhere

****
****

No virus found in this message.
Checked by AVG - www.avg.com
Version: 2015.0.5941 / Virus Database: 4354/9887 - Release Date: 05/28/15

RE: Iqudos software

Salem Hesham

Wed 11/11/2015 15:04

Sent Items

To: Jefferson Kieran (RKB) Consultant Urological Surgeon <Kieran.Jefferson@uhcw.nhs.uk>;

Thanks for your support.

From: Jefferson Kieran (RKB) Consultant Urological Surgeon [Kieran.Jefferson@uhcw.nhs.uk]
Sent: 11 November 2015 15:00
To: Salem Hesham
Subject: RE: Iqudos software

Great

Let me know if you need anything from us

Thanks

Kieran

From: Hesham Salem [mailto:mzxhs@nottingham.ac.uk]
Sent: 11 November 2015 14:53
To: Jefferson Kieran (RKB) Consultant Urological Surgeon
Cc: Muzondo Mandinika (RKB) CNS - Urology
Subject: RE: Iqudos software

Dear Mr Jefferson,

Many thanks for getting back to me. I can use the old data on the Iqudos database which will be completely anonymised by the company. I have been in touch already with your audit and R&D department and they have confirmed that this does not need their approval and suggested that I get in touch directly with your department. i am happy to discuss this further and fill out any relevant forms.

Kind regards,

Hesham

From: Jefferson Kieran (RKB) Consultant Urological Surgeon [Kieran.Jefferson@uhcw.nhs.uk]
Sent: 10 November 2015 14:18
To: Salem Hesham
Cc: Muzondo Mandinika (RKB) CNS - Urology
Subject: Iqudos software

Hi Hesham

Thanks for your enquiry.

We would be very happy to help but have not been using the Iqudos software for a while now as it could never integrate data collection from our databases.

Mandi - can you confirm how long we used it for please?

Thanks

Kieran



DISCLAIMER: This email and any attachments are confidential and may be privileged or otherwise protected from disclosure. If you are not the intended recipient you must not copy this message or any attachment or disclose the contents to any other person, nor take any action in reliance of its contents: to do so is strictly prohibited and may be unlawful. If you have received this message in error please accept our apologies and contact us at the address below or by email at info@uhcw.nhs.uk
Any files attached to this email will have been checked by us with virus detection software before transmission. You should carry out your own virus checks before opening any attachment. UHCW (University Hospitals Coventry and Warwickshire NHS Trust) accepts no liability for any loss or damage which may be caused by software viruses.
Any views or opinions expressed are those of the author and do not represent the views of the Trust unless explicitly stated.
The information contained in this email may be subject to public disclosure under the Freedom of Information Act 2000. Unless the information is legally exempt from disclosure, the confidentiality of this email and your reply cannot be guaranteed.
Address: UHCW NHS Trust, Executive suite, 3rd Floor Rotunda, University Hospital, Clifford Bridge Road, Coventry, CV2 2DX

This message and any attachment are intended solely for the addressee and may contain confidential information. If you have received this message in error, please send it back to me, and immediately delete it.

Please do not use, copy or disclose the information contained in this message or in any attachment. Any views or opinions expressed by the author of this email do not necessarily reflect the views of the University of Nottingham.

This message has been checked for viruses but the contents of an attachment may still contain software viruses which could damage your computer system, you are advised to perform your own checks. Email communications with the University of Nottingham may be monitored as permitted by UK legislation.

Appendix 4: Protocol of the knowledge validation study



Validation of Stable Prostate Cancer knowledge based decision rules

Draft	07 / Final Version 6.0 11/12/14
Short title:	Stable Prostate Cancer knowledge validation
IRAS Project ID:	131767
Study Sponsor:	University of Nottingham
Sponsor reference:	Non-allocated
Funding Source:	Funded by University of Nottingham and Urology department of Royal Derby Hospital

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

STUDY PERSONNEL AND CONTACT DETAILS

Sponsor:	University of Nottingham Contact name Ms Angela Shone Head of Research Governance Research and Graduate Services King's Meadow Campus Lenton Lane Nottingham NG7 2NR
Chief investigator:	Mr Jonathan Lund Associate Professor, University of Nottingham Clinical Sciences Wing, the Medical School, Royal Derby Hospital Uttoxeter Road Derby DE22 3DT Phone:01332 724 694 Email: jon.lund@nottingham.ac.uk
Co-investigator:	Dr Daniele Soria Research Fellow and Lecturer, Faculty of Science, Room B32 Computer Science Jubilee Campus Wollaton Road Nottingham NG8 1BB Phone: 0115 84 68379 Email: daniele.soria@nottingham.ac.uk
Co-investigator and monitor:	Mr Hesham Salem PhD Student, University of Nottingham The Medical School, Royal Derby Hospital Uttoxeter Road Derby DE22 3DT Phone: 07726615815 Email: mzxhs@nottingham.ac.uk
Clinical Study Coordinating Centre:	The Medical School Royal Derby Hospital Uttoxeter Road Derby DE22 3DT
Clinical investigator:	Mr Sadmeet Singh Urology consultant, Royal Derby Hospital, Derby, DE22 3NE Phone: 01332 340 131 Email: sadmeetsingh@nhs.net

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

Clinical investigator:	Mr Richard Stanford Urology consultant, Royal Derby Hospital, Derby, DE22 3NE Phone: 01332 340 131 Email: richard.stanford@nhs.net
Statistical advice	Lynda Cochrane, Director, Clinical Statistics Consultants, Phone: 01382561058 Email: lyndaacochrane@yahoo.com http://clinicalstatisticsconsultants.moonfruit.com/

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

SYNOPSIS

Title	Validation of Stable Prostate Cancer knowledge based decision rules
Short title	SPC knowledge validation
Chief Investigator	Mr Jon Lund
Objectives	The main objective is to validate the acquired knowledge for managing SPC to facilitate further development of decision support tools. All major rule errors will be identified and corrected them to refine the knowledge base. The secondary objective is to estimate the agreement between the rule-based decision system and each of two Urology specialists who are experts in SPC follow up.
Study Configuration	Single tertiary centre database of patients with stable prostate cancer disease
Setting	Cases will be acquired from this existing database of patient seen in a Uro-oncology specialist nurse clinic.
Sample size estimate	100 anonymised cases will be identified from the database and checked against the study inclusion criteria.
Number of investigators	Two Urology specialists who are experts in SPC follow up.
Eligibility criteria	All adult patients with prostate cancer who had their primary treatment assigned by the cancer multidisciplinary team and are known to have stable Prostate Specific Antigen (PSA) levels with no disease or treatment related symptoms.
Description of interventions	The SPC decision tree will be validated and tested in this study to develop a comprehensive and accurate knowledge base to be a platform for further development of decision support tools in including complex Artificial Intelligent systems.
Duration of study	Twenty to twenty two weeks
Methods of analysis	The kappa coefficient will be used to estimate the true agreement of the two clinical investigators with the cases solved by the decision trees while omitting the real outcome. The investigators will rate the decision rules in each case and this will be used to analyse the rules errors. The rules analysis will stratify the rules into three categories: pass, pass with possible updates and fail.

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

ABBREVIATIONS

AI	Artificial Intelligence
CI	Chief Investigator overall
CM	Clinical Monitor
CRF	Case Report Form
DT	Decision trees
GCP	Good Clinical Practice
HWR	Hand written rules
NHS	National Health Service
PI	Principal Investigator at a local centre
PIS	Participant Information Sheet
Pca	Prostate Cancer
PSA	Prostate Specific Antigen
REC	Research Ethics Committee
R&D	Research and Development department
SPC	Stable Prostate Cancer
UoN	University of Nottingham

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

TABLE OF CONTENTS

STUDY PERSONNEL AND CONTACT DETAILS **2**

SYNOPSIS **4**

ABBREVIATIONS **5**

STUDY BACKGROUND INFORMATION AND RATIONALE **8**

STUDY OBJECTIVES AND PURPOSE **8**

PURPOSE 8
PRIMARY OBJECTIVE 9
SECONDARY OBJECTIVES 9

STUDY FLOW CHART **9**

STUDY DESIGN **10**

STUDY CONFIGURATION 10
ANONYMISATION AND BLINDING 10
STUDY MANAGEMENT 10
DURATION OF THE STUDY 10
End of the Study 10
SELECTION OF CASES 11
Eligibility criteria 11
Inclusion criteria 11
Exclusion criteria 11
Informed consent 11
STUDY REGIMEN 11
Compliance 12
Criteria for terminating the study 12

ANALYSES **12**

Methods (Study flow chart) 12
Sample size and justification 12

ADVERSE EVENTS **13**

ETHICAL AND REGULATORY ASPECTS **13**

ETHICS COMMITTEE AND REGULATORY APPROVALS 13
INFORMED CONSENT AND PARTICIPANT INFORMATION 13
RECORDS 13
Case Report Form (Appendix 1) 13
Source documents 14
Direct access to source data / documents 14
DATA PROTECTION 14

QUALITY ASSURANCE & AUDIT **14**

INSURANCE AND INDEMNITY 14
STUDY CONDUCT 14
STUDY DATA 14
RECORD RETENTION AND ARCHIVING 14
DISCONTINUATION OF THE STUDY BY THE SPONSOR 15
STATEMENT OF CONFIDENTIALITY 15

PUBLICATION AND DISSEMINATION POLICY **15**

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

USER AND PUBLIC INVOLVEMENT **15**
STUDY FINANCES **15**
Funding source 15
Participant stipends and payments 15
SIGNATURE PAGES **16**
REFERENCES **17**
APPENDIX ONE **18**
APPENDIX TWO **20**
APPENDIX THREE **22**
APPENDIX FOUR **24**

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

STUDY BACKGROUND INFORMATION AND RATIONALE

Prostatic adenocarcinoma is the most common cancer to be diagnosed in men (25% of all cancers in 2010) with an estimated incidence of 104 cases per 100,000 in 2010 [1]. It has been predicted that this rate could increase seven-fold by 2030 [2] due to the advancement of the disease diagnostic tools and increase of average population life expectancy.

In the UK, the initial decision to treat is usually established in the local or regional Urology cancer multidisciplinary meeting where the treatment can be categorised into disease cure, such as surgery, high dose radiotherapy etc. or disease control treatment including palliation. Following the initiation of treatment and disease stabilisation, the patients can be followed up every 6 to 12 months to looks for early signs of progression and complications of treatment.

It has been recommended by the National Institute of Health Care Excellence (NICE) to provide the follow up care for these patients with stable disease in the community [3, 4] and It is not clear from the literature if this national recommendation has been widely accepted by clinicians in the NHS. In fact, a recent qualitative questionnaire based review [5] demonstrated a palpable concern among various health care workers on the implementation of this recommendation without any existing measures to support the decision making in the primary care in the form of personnel training, direct access to MDT and decision support tools.

The developments of such supportive tools share a common initial step of acquiring accurate and valid knowledge from the domain experts, which can then be constructed into clinical pathways, decision trees and AI (artificial intelligence) based systems etc. The later computer based systems face more significant challenges in this stage since it is usually orchestrated by a knowledge engineer who is not by default a domain expert or familiar with the knowledge and its natural language. The knowledge engineer faces more obstacles while representing the elicited knowledge fragments to the system developers in a specific model framework. This requires deep understanding of the elicited knowledge fragments and the reasoning scheme used by the experts to solve the problem or perform the specified task. This is in addition to setting the knowledge hierarchies, for instance if a patient has renal failure and bone cancer metastasis, which one is considered more urgent in the decision-making? By the end of this complex process, the knowledge engineer should be able to express the domain knowledge in its three types as factual, inferential and strategic knowledge.

This knowledge base development stage is considered the bottleneck for developing these systems and can arrest its progress [6-8]. A further more challenge to these AI based system development is the 'knowledge validation', where they lack conformity with no standard model to recommend when the validation should commence and what methodology to be followed. Interestingly the hand written integrated care pathways face a similar limitation where they lack formality in validation while the existing validation tools focusing on the pathway development rather than the patients' outcome [9].

STUDY OBJECTIVES AND PURPOSE

PURPOSE

Developing a comprehensive, valid and accurate knowledge base for managing SPC in a format defining the three knowledge types (factual, inferential and strategic) in order to facilitate further development of decision support tools in this domain.

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

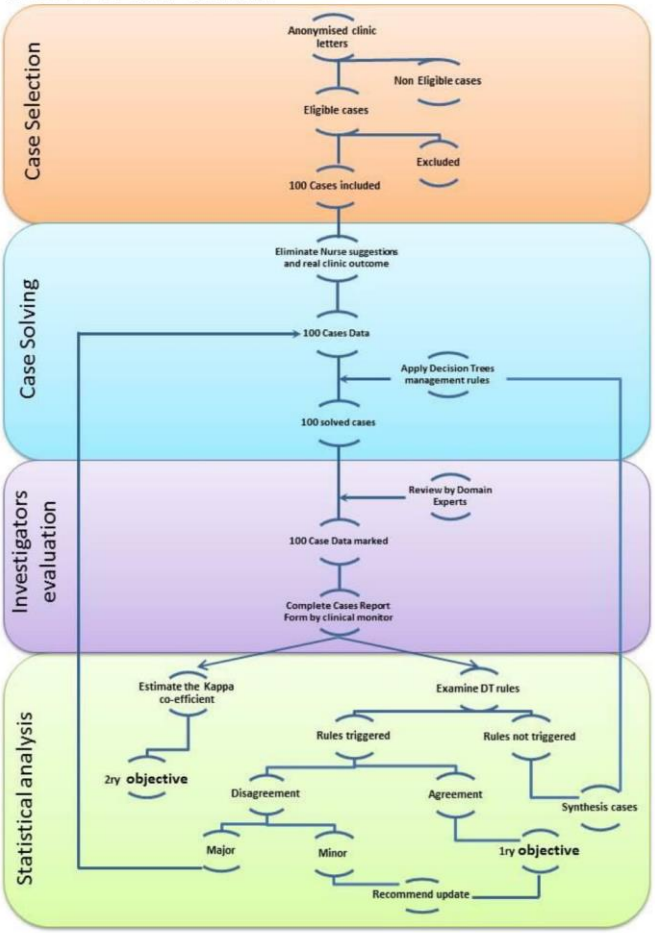
PRIMARY OBJECTIVE

The main objective is to validate the acquired knowledge by estimating the true experts' agreement with the decision rules suggestions for real patients presentation gathered retrospectively from a hospital's database. Inaccurate rules will be identified and revised.

SECONDARY OBJECTIVES

The secondary objective is to estimate the agreement between the DT & HWR and each of two experts.

STUDY FLOW CHART



This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

STUDY DESIGN

STUDY CONFIGURATION

Real, anonymised data will be collected retrospectively from an existing database of patients reviewed in a specialist nurse-led clinic. The clinical input data from each patient visit will be reasoned by the DT & HWR and presented, in the form of a case presentation and a suggested outcome mapped to our study decision trees, to two independent assessors with comparable experience and expertise (Appendix 1: Sample of a case report form). Each rater's agreement with the DT & HWR will be estimated.

ANONYMISATION AND BLINDING

All patient identifiers will be concealed and known only to the hospital database officer. Each case will be assigned a unique study identification number, which will be maintained by the hospital database officer for future data auditing where necessary.

The study CM will omit the real specialist nurse suggestions for managing the presenting case and instead synthesise suggestions according to the decision trees to have a study 'solved case' (Appendix 1). Thus each expert will be blind to the actual patient outcome and the other investigator's decision.

STUDY MANAGEMENT

The Chief Investigator has overall responsibility for the study and shall oversee all study management.

DURATION OF THE STUDY

The total duration is predicted to be **eighteen to twenty** weeks.
The anonymised clinic letter will be examined by the study CM in order to select the cases meeting the inclusion criteria and providing the diversity required to test all of the decision rules in the DT. It is predicted to take **two weeks** pending the cases' availability.
The included cases will be solved by reasoning through the knowledge in the DT before being presented to the clinical investigators. This can be processed in **two weeks**.
Each of the three investigators will examine the decision rule in one hundred cases over **twelve weeks**.
The last stage of study will include the statistical analysis which will require **two to four weeks**.
Participant Duration: Not applicable as we are not recruiting any participants.

End of the Study

The study will be terminated after correcting any major errors in the decision rules (please see study flow chart).

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

SELECTION OF CASES

Cases will be selected from the existing database, randomly in the first instance. The CM may exclude cases where there is a repetition (i.e. triggering the same pathway in the decision tree) and request more cases from the hospital data base officer if necessary. Some of the decision rules are formulated for rare clinical scenarios, which may not be represented in the database. The CM will simulate such rules by altering a number of real cases. These synthetic records will be reviewed by an independent expert before being signed off by the chief investigator for inclusion in the study. In total, one hundred cases will be selected.

Eligibility criteria

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

Inclusion criteria

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

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

With or without (optional requirement):

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

Exclusion criteria

- Cases with newly diagnosed prostate cancer that are waiting for treatment decision.
- Cases with known prostate cancer that are being evaluated for known recent disease progression (i.e. not SPC).
- Cases with no clinical details of their initial cancer presentation such as Gleason score, PSA, and tumour grade which would influence the follow up decision making.

Informed consent

NA

STUDY REGIMEN

The CM will follow the methodology mentioned in the “Selection of cases” section above to select 100 cases while considering various clinical presentations triggering all of the decision trees pathways (Appendix 2).

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

Following the case selection, the CM will eliminate the nurse's recommendations and real clinic outcome to replace it by suggestions derived from the decision trees (Appendix 1). Case-solving according to the decision tree is an objective process that can only lead to one of four main outcomes (*Routine follow up*, *Urgent follow up*, *Routine follow up with symptom specific advice* and *Refer to the hospital*) and other disease management suggestions. Each solved case will be presented as in the sample CRF (Appendix 1) with a space for the investigator to mark as A (agree) or D (disagree) for the main outcome and each of the other suggestions. There will be a free text box for gathering additional information. In cases of disagreement, the investigator will document the category of the rule error (minor or major).

Data will be recorded by the CM in the study Data Report Form (Appendix 3: Sample DRF) and submitted for analysis.

Compliance

NA

Criteria for terminating the study

There are no expected adverse effects of the study to warrant its termination.

ANALYSES

Methods (Study flow chart)

All cases of disagreement between an expert and DT & HWR will be identified and investigated. We will aim to identify the tested and untested decision rule in addition to the individual rule errors. The data from the CRF will be documented in the Rules Report Form (RRF, Appendix 4); a report on each tested rule will be produced.

For the tested rules, each will be individually examined for success or failure based on the investigators comments analysis, number reported as minor errors and number reported as major errors before being categorised as *Pass* (all agreement and no negative comments), *Pass with possible updates* (if only minor disagreements reported) and *Fail* (major disagreement and/or negative comments).

Failing rules will be updated and peer reviewed with the clinical investigators before further testing.

To meet the secondary objective, the weighted Kappa statistic will be calculated (Estimate, 95% Confidence Interval) to summarise the level of agreement between the DT & HWR each rater.

Sample size and justification

The objective of this study is not to make statistical inferences so no sample size / power calculations were carried out. A sample size of 100 is a realistic to achieve in the time-scale and will enable us to explore all the rules.

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

ADVERSE EVENTS

The data will be collected retrospectively from an existing database and the study is not expected to cause any adverse events to the patients. Thus adverse events data will not be collected.

ETHICAL AND REGULATORY ASPECTS

The first risk identified is breaching patient's confidentiality, which will be eliminated by anonymising all patients' identifiers before submitting the cases to the CM.

The second risk is clinical investigator rating bias, which will be eliminated by selecting reviewers who did not contribute in the development of the decision rules and are not known to have any conflicting interest in the outcome of this study.

This study is not expected to have any direct or indirect effects on any patient or clinical investigator and all the data will be retrospectively extracted from an existing hospital database.

ETHICS COMMITTEE AND REGULATORY APPROVALS

This study will only analyse anonymised data with no effect on patients' care. Thus REC approval was deemed unnecessary.

INFORMED CONSENT AND PARTICIPANT INFORMATION

NA

RECORDS

Case Report Form (Appendix 1)

The data will be presented to the investigators in the form of a clinical letter as in sample CRF (Appendix 1) that would have a unique study number and information on:

- Patient's age
- Diagnosis
- Gleason Score (if available)
- Metastasis (if present)
- Previous and active cancer treatment
- PSA at presentation
- PSA nadir (if available)
- PSA most recent (if available)
- Symptoms (if present)
- Other blood test results (if available)
- Recommendation according to the DT with a space for the investigator to rate each recommendation and enter own comment where appropriate.

Data Report Form: (Appendix 2)
This will be completed by the CM to report the data from the CRF.

Rules Report Form: (Appendix 4)
The CM will fill this form to report the data per each of the 36 decision rules.

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

Source documents

Source documents shall be filed at the investigator's site and only study staff shall have access to study documentation other than the regulatory requirements listed below.

Direct access to source data / documents

The CRF and all source documents shall made be available at all times for review by the Chief Investigator, Sponsor's designee and inspection by relevant regulatory authorities.

DATA PROTECTION

Only anonymised data is included in this study and the data protection act would not apply.

QUALITY ASSURANCE & AUDIT

INSURANCE AND INDEMNITY

Insurance and indemnity for clinical study participants and study staff is covered within the NHS Indemnity Arrangements for clinical negligence claims in the NHS, issued under cover of HSG (96)48. There are no special compensation arrangements, but study participants may have recourse through the NHS complaints procedures.

The University of Nottingham as research Sponsor indemnifies its staff, research participants and research protocols with both public liability insurance and clinical trials insurance. These policies include provision for indemnity in the event of a successful litigious claim for proven non-negligent harm.

STUDY CONDUCT

The Clinical monitor shall carry out a site systems audit and an audit report shall be made.

STUDY DATA

The CM shall carry out monitoring of study data as an ongoing activity.

Entries on CRFs will be verified by inspection against the source data. A sample of CRFs (10% or as per the study risk assessment) will be checked on a regular basis for verification of all entries made. In addition, the subsequent capture of the data on the study database will be checked. Where corrections are required, these will carry a full audit trail and justification.

Study data, and evidence of monitoring and systems audits, will be made available for inspection by the REC as required.

RECORD RETENTION AND ARCHIVING

In compliance with the ICH/GCP guidelines, regulations and in accordance with the University of Nottingham Code of Research Conduct and Research Ethics, the Chief or local Principal Investigator will maintain all records and documents regarding the conduct of the study. These will be retained for at least 7 years or for longer if required. If the

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

responsible investigator is no longer able to maintain the study records, a second person will be nominated to take over this responsibility.

The study documents held by the Chief Investigator on behalf of the Sponsor shall be finally archived at secure archive facilities at the University of Nottingham. This archive shall include all anonymised audio recordings, study databases and associated meta-data encryption codes.

DISCONTINUATION OF THE STUDY BY THE SPONSOR

The Sponsor reserves the right to discontinue this study at any time for failure to meet expected enrolment goals, for safety or any other administrative reasons. The Sponsor shall take advice as appropriate in making this decision.

STATEMENT OF CONFIDENTIALITY

Data generated as a result of this study will be available for inspection on request by the participating physicians, the University of Nottingham representatives, the REC, local R&D Departments and the regulatory authorities.

PUBLICATION AND DISSEMINATION POLICY

The results will be submitted for peer reviewed publication in relevant scientific journal and specialty academic meetings.

USER AND PUBLIC INVOLVEMENT

NA

STUDY FINANCES

Funding source

This study is not receiving any external funding.

Participant stipends and payments

NA

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

SIGNATURE PAGES

Signatories to Protocol:

Chief investigator: Mr Jonathan Lund

Signature:

Date:

Co-investigator: Dr Daniele Soria

Signature:

Date:

Co-investigator and monitor: Mr Hesham Salem

Signature:

Date:

Clinical investigator: Mr Sadmeet Singh

Signature:

Date:

Clinical investigator: Mr Richard Stanford

Signature:

Date:

Statistical advice Lynda Cochrane

Signature:

Date:

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

REFERENCES

1. UK, C.R., *Prostate Cancer (C61), Average Number of New Cases per Year and Age-Specific Incidence Rates, Males, UK, 2008-2010*. 2012.
2. Mistry, M., et al., *Cancer incidence in the United Kingdom: projections to the year 2030*. British Journal of Cancer, 2011. **105**(11): p. 1795-1803.
3. NICE, *Prostate cancer: diagnosis and treatment CG28*. 2008.
4. NICE, *Prostate cancer: diagnosis and treatment CG175*. 2014, National Institute for Health and Care Excellence.
5. Watson, E.K., et al., *Views of health professionals on the role of primary care in the follow-up of men with prostate cancer.* Family Practice, 2011. **28**(6): p. 647-654.
6. Achour, S.L., et al., *Application of information technology - A UMLS-based knowledge acquisition tool for rule-based clinical decision support system development.* Journal of the American Medical Informatics Association, 2001. **8**(4): p. 351-360.
7. Benbasat, I. and J.S. Dhaliwal, *A framework for the validation of knowledge acquisition.* Knowledge Acquisition, 1989. **1**(2): p. 215-233.
8. Pandey, B. and R.B. Mishra, *Knowledge and intelligent computing system in medicine.* Computers in Biology and Medicine, 2009. **39**(3): p. 215-230.
9. Vanhaecht, K., et al., *Clinical pathway audit tools: a systematic review.* Journal of nursing management, 2006. **14**(7): p. 529-37.

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

APPENDIX ONE

Department of Urology
Prostate Cancer Telephone Clinic
NHS Number: 0123456789
Date of Clinic: 15/11/2010
Consultant: MJH

Derby Hospitals NHS
NHS Foundation Trust

Royal Derby Hospital
Uttoxeter Road
Derby
DE22 3NE

Tel: 01332 340131
Minicom: 01332 254944
contactus@derbyhospitals.nhs.uk
www.derbyhospitals.nhs.uk

PRIVATE AND CONFIDENTIAL
Dr. D SPOCK
The Old Church Surgery
17 Bishop Street South
Pendle
Gifford
Derby

Dear Dr. D SPOCK

Stable Prostate Cancer Pathway: **Radical Prostatectomy**

Many thanks for allowing Mr. Test Subject: 643 to be reviewed by the stable prostate cancer clinic. Mr. Test Subject: 643 is 68 years old and was diagnosed with Gleason score 7 cancer on 25 February 2008. His TMN stage was recorded as cT2a pT2 N0 M0.

His recorded PSA at presentation was 7.80 ng/ml.

Treatments:
18 Apr 2008 Active Surveillance
31 Mar 2009 Radical Retropubic Prostatectomy

Today's Findings

The level of PSA present in the last available blood test is 0.00 ng/ml.

Erectile Dysfunction.

Clinical Management

The patient is being monitored on the radical prostatectomy management pathway and his latest PSA result is showing a normal un-measurable reading. Thus, an appointment will be made to see the patient again in 6 months' time (Rule x)

Mr. Test Subject: 643 is concerned regarding erectile dysfunction. It would be worth trying him with a PDE inhibitor such as sildenafil. If there are contraindications or the patient prefers, an alternative would be the use of a vacuum tumescence device. If the patient requires injection therapy with prostaglandin analogues then referral to secondary care may be needed. (Rule y)

Blood Test Name	Unit	09/11/2010
Alanine Transaminase (ALT)	s lu/L	4.6
Alkaline Phosphatase (ALK)	lu/L	103
Corrected Calcium	mmol/L	2.3
Creatinine	umol/L	49
Haemoglobin (Hb)	g/dl	17.5
Potassium	mmol/L	3.6
Prostate Specific Antigen (PSA)	ug/L	0

1 | Page

London Road Community Hospital,
London Road, Derby. DE1 2QY
Smoking is not permitted anywhere in the buildings and grounds of Derby's Hospitals. For advice and support about giving up smoking please call Free Phone 0800 707 6870

Chief Executive: Susan James
Chair: John Rivers

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

Department of Urology
Prostate Cancer Telephone Clinic
NHS Number: 0123456789
Date of Clinic: 15/11/2010
Consultant: MJH

Derby Hospitals NHS
NHS Foundation Trust

Royal Derby Hospital
Uttoxeter Road
Derby
DE22 3NE

Tel: 01332 340131
Minicom: 01332 254944
contactus@derbyhospitals.nhs.uk
www.derbyhospitals.nhs.uk

Kind Regards,

Nurse Sharon Williams
On behalf of the stable prostate cancer clinic

Rule number	Agree	Disagree	Alternative if disagree	Comments(MINOR OR MAJOR ERROR)
X			1	
			2	
			3	
			4	
Y			1	
			2	
			3	
			4	

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

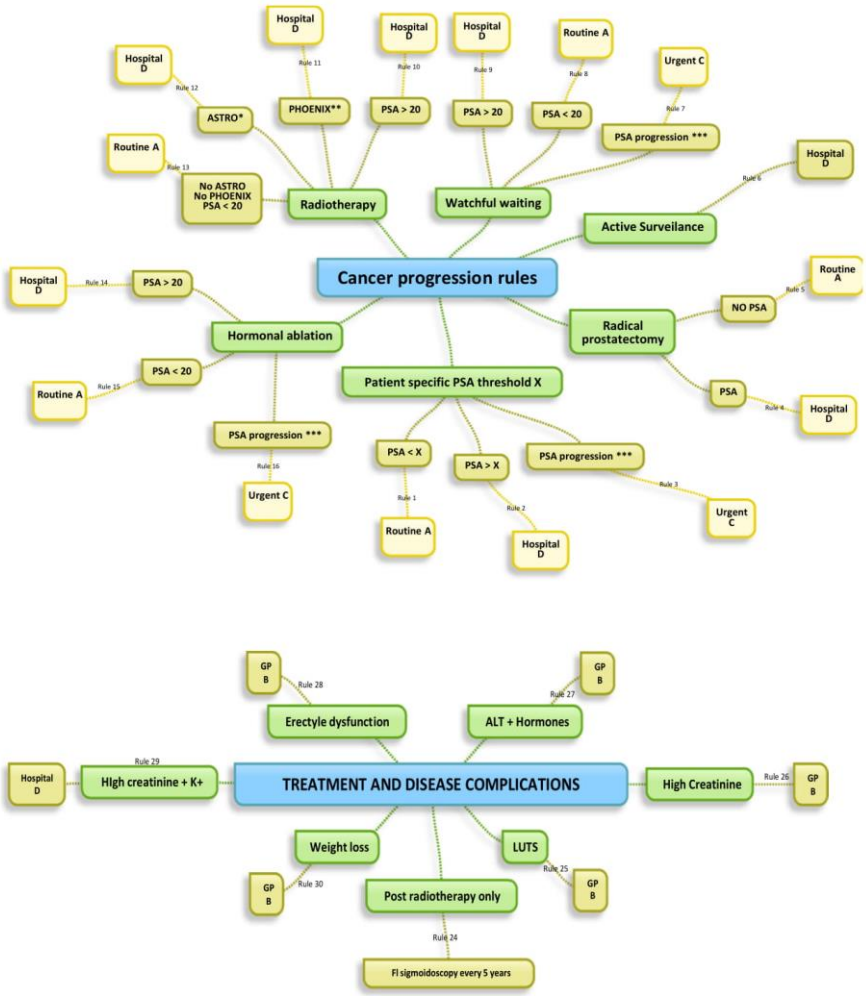
Name	Investigator 1
Sign	
Date	

London Road Community Hospital,
London Road, Derby. DE1 2QY
Smoking is not permitted anywhere in the buildings and grounds of Derby's Hospitals. For advice and support about giving up smoking please call Free Phone 0800 707 6870

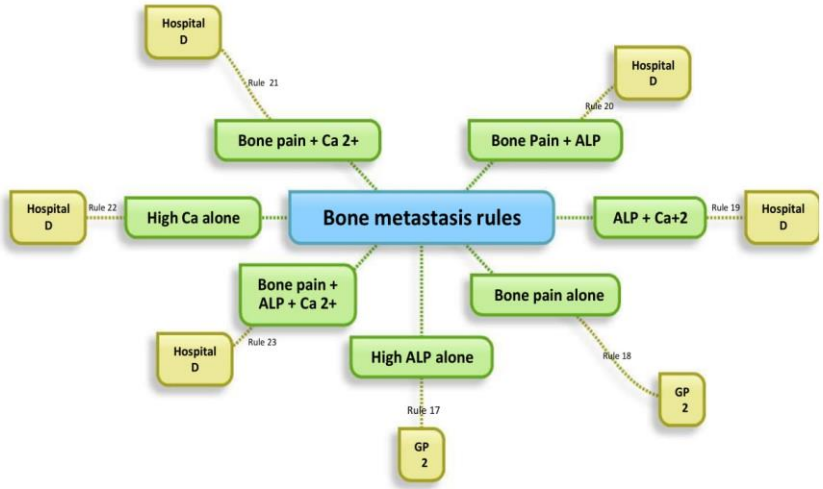
Chief Executive: Susan James
Chair: John Rivers

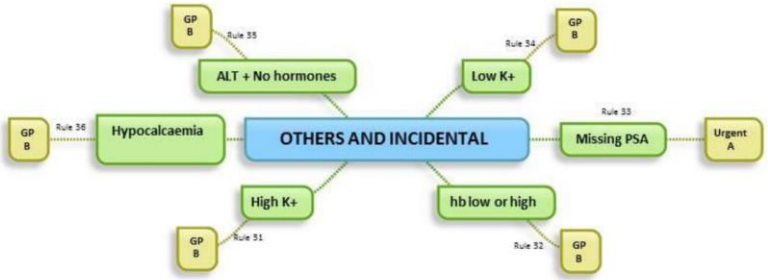
This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

APPENDIX TWO



This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham





This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

APPENDIX THREE

Data Report Form (DRF)

Name of the investigator:
Place of the investigation: UoN
Study Identification No.:
Study Duration:
DRF Identification number:

Case number	Rule number	DT outcome	Investigator outcome

SPC knowledge validation DRF v 2.0 1

This form is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

CLINICAL MONITOR
SIGN: Date ☐☐ / ☐☐☐ / ☐☐
PRINT NAME: HESHAM SALEM

This form is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

APPENDIX FOUR

Rules Report Form

Study Identification No:
Place of the investigation: UoN
Study Duration:
DRF included:

Rule n.	n used	n. Agree	n. Disagree	Minor errors	Major errors	Comments	Pass or Fail

SPC knowledge validation RRF v 0 .1 1

This form is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

CLINICAL MONITOR	
SIGN:	Date ☐☐ / ☐☐☐ / ☐☐
PRINT NAME: HESHAM SALEM	

This form is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

Appendix 5: Clinical cases presented to investigators for rating in the knowledge validation study

Department of Urology
Prostate Cancer Telephone Clinic
NHS Number: 0123456789
Date of Clinic: 15/11/2010
Consultant: MJH

Derby Hospitals NHS
NHS Foundation Trust

Royal Derby Hospital
Uttoxeter Road
Derby
DE22 3NE

Tel: 01332 340131
Minicom: 01332 254944
contactus@derbyhospitals.nhs.uk
www.derbyhospitals.nhs.uk

PRIVATE AND CONFIDENTIAL
Dr. D SPOCK
The Old Church Surgery
17 Bishop Street South
Pendle
Gifford
Derby

Dear Dr. D SPOCK

Stable Prostate Cancer Pathway: **Radical Prostatectomy**

Many thanks for allowing Mr. Test Subject: 643 to be reviewed by the stable prostate cancer clinic. Mr. Test Subject: 643 is 68 years old and was diagnosed with Gleason score 7 cancer on 25 February 2008. His TMN stage was recorded as cT2a pT2 N0 M0.

His recorded PSA at presentation was 7.80 ng/ml.

Treatments:
18 Apr 2008 Active Surveillance
31 Mar 2009 Radical Retropubic Prostatectomy

Today's Findings

The level of PSA present in the last available blood test is 0.00 ng/ml.

Erectile Dysfunction.

Clinical Management

The patient is being monitored on the radical prostatectomy management pathway and his latest PSA result is showing a normal un-measurable reading. Thus, an appointment will be made to see the patient again in 6 months' time (Rule x)

Mr. Test Subject: 643 is concerned regarding erectile dysfunction. It would be worth trying him with a PDE inhibitor such as sildenafil. If there are contraindications or the patient prefers, an alternative would be the use of a vacuum tumescence device. If the patient requires injection therapy with prostaglandin analogues then referral to secondary care may be needed. (Rule y)

Blood Test Name	Unit	09/11/2010
Alanine Transaminase (ALT)	s lu/L	4.6
Alkaline Phosphatase (ALK)	lu/L	103
Corrected Calcium	mmol/L	2.3
Creatinine	umol/L	49
Haemoglobin (Hb)	g/dl	17.5
Potassium	mmol/L	3.6
Prostate Specific Antigen (PSA)	ug/L	0

Rule Applied to case specified clinical presentation

London Road Community Hospital,
London Road, Derby. DE1 2QY

Smoking is not permitted anywhere in the buildings and grounds of Derby's Hospitals. For advice and support about giving up smoking please call Free Phone 0800 707 6870

Chief Executive: Susan James
Chair: John Rivers

Department of Urology
Prostate Cancer Telephone Clinic
NHS Number: 0123456789
Date of Clinic: 15/11/2010
Consultant: MJH

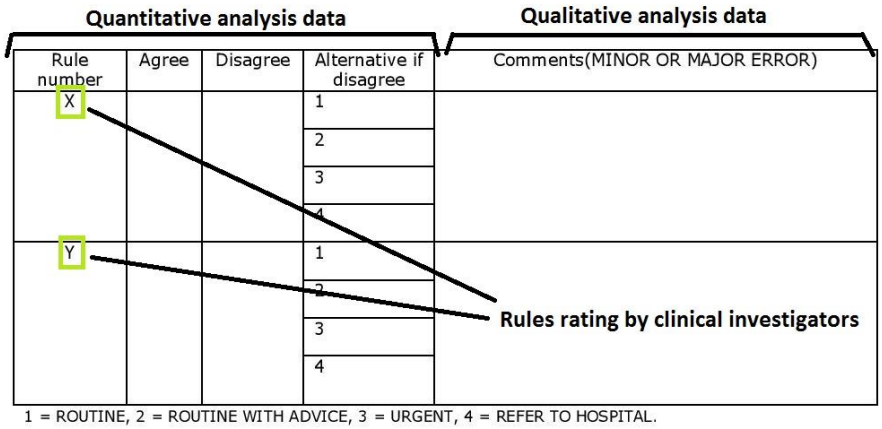
Derby Hospitals NHS
NHS Foundation Trust

Royal Derby Hospital
Uttoxeter Road
Derby
DE22 3NE

Tel: 01332 340131
Minicom: 01332 254944
contactus@derbyhospitals.nhs.uk
www.derbyhospitals.nhs.uk

Kind Regards,

Nurse Sharon Williams
On behalf of the stable prostate cancer clinic



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

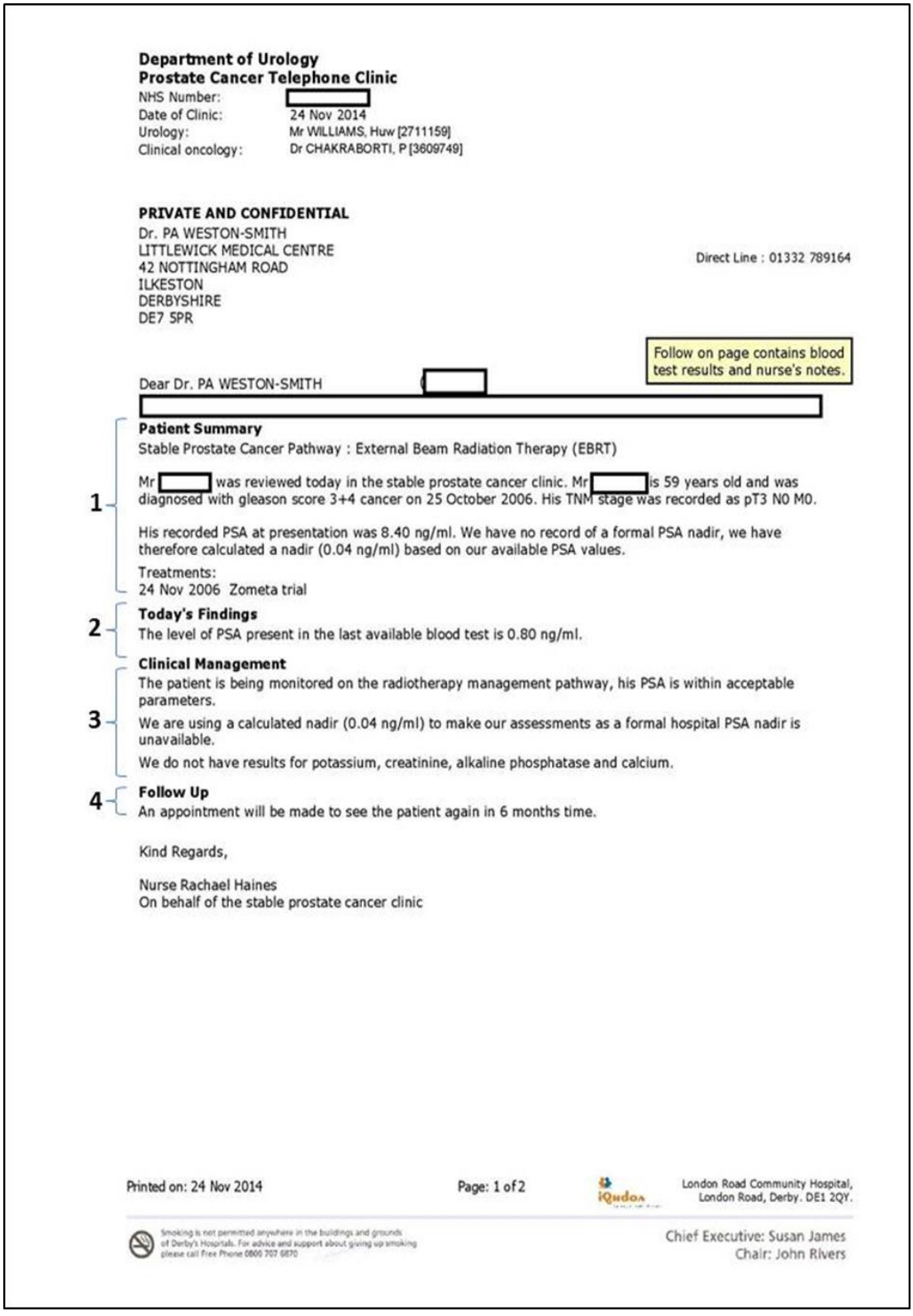
Name	Investigator 1
Sign	
Date	

London Road Community Hospital,
London Road, Derby. DE1 2QY

Smoking is not permitted anywhere in the buildings and grounds of Derby's Hospitals. For advice and support about giving up smoking please call Free Phone 0800 707 6870

Chief Executive: Susan James
Chair: John Rivers

Appendix 6: Electronic letter generated by the CDSS in the live validation study



Department of Urology
Prostate Cancer Telephone Clinic
NHS Number: [redacted]
Date of Clinic: 24 Nov 2014
Urology: Mr WILLIAMS, Huw [2711159]
Clinical oncology: Dr CHAKRABORTI, P [3609749]

PRIVATE AND CONFIDENTIAL
Dr. PA WESTON-SMITH
LITTLEWICK MEDICAL CENTRE
42 NOTTINGHAM ROAD
ILKESTON
DERBYSHIRE
DE7 5PR

Direct Line : 01332 789164

Follow on page contains blood test results and nurse's notes.

Dear Dr. PA WESTON-SMITH [redacted]

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

1 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.

Treatments:
24 Nov 2006 Zometa trial

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

Clinical Management
The patient is being monitored on the radiotherapy management pathway, his PSA is within acceptable parameters.

3 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 Rachael Haines
On behalf of the stable prostate cancer clinic

Printed on: 24 Nov 2014 Page: 1 of 2 iQudos London Road Community Hospital, London Road, Derby. DE1 2QY.

Smoking is not permitted anywhere in the buildings and grounds of Derby's Hospitals. For advice and support about giving up smoking please call Free Phone 0800 707 6870

Chief Executive: Susan James
Chair: John Rivers

[redacted]

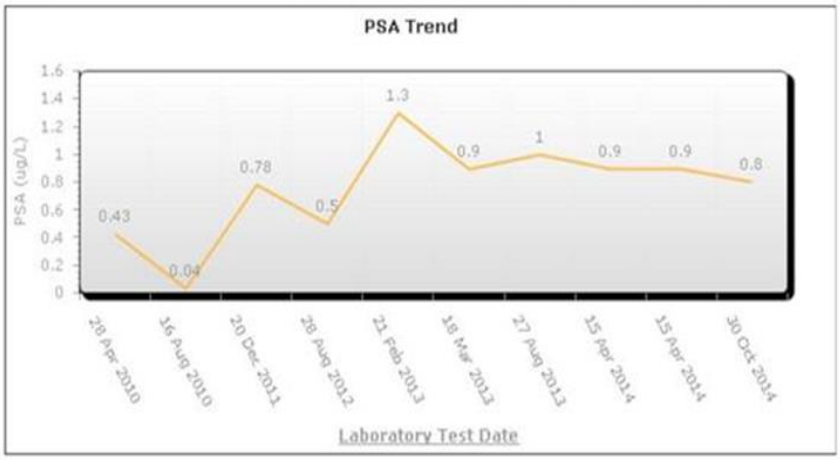
Nurse's Notes:

5 Mr [redacted] is aware of his recent psa. He has no bothersome LUTS, has recently lost weight but this is planned weight loss. He is happy to remain on the 6 monthly telephone clinic.

Recent appointments with associated blood test results:

Blood Test Name	Units	3/5/2013	4/26/2014	11/24/2014
Prostate Specific Antigen (PSA)	ug/L	1.3 (21 Feb 2013)	0.9 (15 Apr 2014)	0.8 (30 Oct 2014)

Key: [-] denotes result not recorded (Date) in parenthesis denotes the lab test date



Smoking is not permitted anywhere in the buildings and grounds of Derby's Hospitals. For advice and support about giving up smoking please call Free Phone 0800 707 6870

Chief Executive: Susan James
Chair: John Rivers

Appendix 7: Example of cases generated for the live validation

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

Derby Hospitals NHS
NHS Foundation Trust

Royal Derby Hospital
Uttoxeter Road
Derby
DE22 3NE

Tel: 01332 340131
Minicom: 01332 254944
contactus@derbyhospitals.nhs.uk
www.derbyhospitals.nhs.uk

Re: Mr Study case DB268

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

Mr DB268 was reviewed 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 pT2c N0 M0.

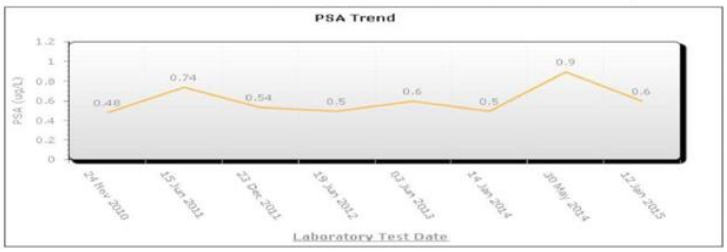
His recorded PSA at presentation was 7.90 ng/ml. His PSA nadir is [0.08 ng/ml].

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	

1 | Page

Appendix 8: Protocol of the live validation

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

A multi-centre prospective validation of an Expert system led follow up service for patients with stable prostate cancer

Draft 1.3 / Final Version 1.0
17th March

Short title:	Validation of an Expert system in stable prostate cancer
IRAS Project ID:	131767
Trial Sponsor:	University of Nottingham
Funding Source:	Derby Teaching Hospitals NHS Foundation Trust

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

TRIAL / STUDY PERSONNEL AND CONTACT DETAILS

Sponsor:	University of Nottingham Contact name Ms Angela Shone Head of Research Governance Research and Graduate Services King's Meadow Campus Lenton Lane Nottingham NG7 2NR Email: sponsor@nottingham.ac.uk
Chief investigator:	Mr Jonathan Lund Associate Professor, University of Nottingham Clinical Sciences Wing, the Medical School, Royal Derby Hospital Uttoxeter Road Derby DE22 3DT Phone: 01332 724 694 Email: jon.lund@nottingham.ac.uk
Co-investigator:	Dr Daniele Soria Research Fellow and Lecturer, Faculty of Science, Room B32 Computer Science Jubilee Campus Wollaton Road Nottingham NG8 1BB Phone: 0115 84 68379 Email: daniele.soria@nottingham.ac.uk
Co-investigator and monitor:	Mr Hesham Salem PhD Student, University of Nottingham The Medical School, Royal Derby Hospital Uttoxeter Road Derby DE22 3DT Phone: 07726615815 Email: mzxhs@nottingham.ac.uk
Clinical Trial Coordinating Centre:	The Urology Department Royal Derby Hospital Uttoxeter Road Derby DE22 3NE

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

Clinical investigator:	Mr Giacomo Caddeo Urology consultant, Royal Derby Hospital, Derby, DE22 3NE Phone: 01332 340 131 Email: giacomo.caddeo@nhs.net
Clinical investigator:	Mr Jon McFarlane Urology Consultant Royal United Hospitals Bath NHS Foundation Trust Combe Park, Bath, Avon, Somerset BA1 3NG Phone: 0789 422 0840 Email: jmcfarlane2@nhs.net
Statistical advice	Dr Lynda Cochrane, Director, Clinical Statistics Consultants, Phone: 01382 561058 Email: Lynda.Cochrane@sky.com http://clinicalstatisticsconsultants.moonfruit.com/

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

SYNOPSIS

Title	A multi-centre prospective verification and validation of an Expert system led follow up service for patients with stable prostate cancer
Short title	Validation of an Expert system in stable prostate cancer
Chief Investigator	Mr Jonathan Lund, Associate Professor, University of Nottingham
Objectives	To validate the Expert system with real cases against two blinded Urology consultants. To estimate the system reliability and repeatability by checking the outcome of similar cases.
Setting	Multicentre anonymised data collection of patients with stable prostate cancer disease presenting to established follow up service.
Sample size estimate	The sample size required to test the null hypothesis that the agreement between the ES and the human domain experts is slight to fair (k_0= 0.4) against the alternative that it is better than moderate (k_1 > 0.61) will be determined when the frequencies of outcomes in each category is known.
Number of Cases	200 (number will vary according to the level of agreement between the clinicians and the Expert System)
Eligibility criteria	All adult patients seen in these clinics with known prostate cancer whom had their primary treatment assigned and delivered with enough information on their PSA levels.
Duration of study	3 months
Randomisation and blinding	The clinical investigators will individually synthesis their own decisions and outcomes for the included cases while being blinded from the Expert system's decisions.
Outcome measures	Kappa statistic
Statistical methods	The primary measure of agreement will the Kappa statistic, estimated using SPSS v 22.

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

ABBREVIATIONS

AE	Adverse Event
CDSS	Clinical Decision Support System
CI	Chief Investigator overall
CM	Clinical Monitor
CNS	Cancer Nurse Specialist
CRF	Case Report Form
DMC	Data Monitoring Committee
ES	Expert System
GCP	Good Clinical Practice
HIS	Hospital Information System
ICF	Informed Consent Form
IG SoC	Information Governance Statement of Compliance
MHRA	Medicine and Healthcare products Regulatory Agency
NHS	National Health Service
NICE	National Institute of Health Care Excellence
NPCA	National Prostate Cancer Audit
PI	Principal Investigator at a local centre
PIS	Participant Information Sheet
Pca	Prostate Cancer
REC	Research Ethics Committee
R&D	Research and Development department
SAE	Serious Adverse Event
SPC	Stable prostate cancer
UK	The United Kingdom

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

TABLE OF CONTENTS

TRIAL / STUDY PERSONNEL AND CONTACT DETAILS **2**

SYNOPSIS **4**

ABBREVIATIONS **5**

TRIAL / STUDY BACKGROUND INFORMATION AND RATIONALE **8**

TRIAL / STUDY OBJECTIVES AND PURPOSE **9**

PURPOSE 9
PRIMARY OBJECTIVE 9
SECONDARY OBJECTIVES 9

EXPERT SYSTEM STRUCTURE **9**

DESCRIPTION OF THE PROSTATE CANCER SYSTEM 10
SYSTEM SECURITY AND FUNCTIONAL ANALYSIS 10

TRIAL / STUDY DESIGN **10**

TRIAL / STUDY CONFIGURATION 10
Primary endpoint 11
Secondary endpoint 11

TRIAL FLOW CHART **12**

Complications from applying the ES in the CNS clinic 13
Stopping rules and discontinuation 13
BLINDING 13
TRIAL/STUDY MANAGEMENT 13
DURATION OF THE TRIAL 13
End of the Trial 13
SELECTION OF CASES 13
Eligibility criteria 13
Inclusion criteria 13
Exclusion criteria 14
Informed consent 14

STATISTICS **14**

Methods 14
Sample size and justification 14

ETHICAL AND REGULATORY ASPECTS **14**

ETHICS COMMITTEE AND REGULATORY APPROVALS 14
RECORDS 15
Case Report Forms 15
Data Report Form 15
Source documents 15
Direct access to source data / documents 15
DATA PROTECTION 15

QUALITY ASSURANCE & AUDIT **15**

INSURANCE AND INDEMNITY 15

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

TRIAL CONDUCT 16
TRIAL DATA 16
RECORD RETENTION AND ARCHIVING 16
DISCONTINUATION OF THE TRIAL BY THE SPONSOR 16
STATEMENT OF CONFIDENTIALITY 16
PUBLICATION AND DISSEMINATION POLICY 16
STUDY FINANCES 16
Funding source 16
SIGNATURE PAGES 17
REFERENCES 18
APPENDIX ONE 19
APPENDIX TWO 21
APPENDIX THREE 22
APPENDIX FOUR 23

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

TRIAL / STUDY BACKGROUND INFORMATION AND RATIONALE

Prostate cancer (Pca) is the most common malignant disease in men and the third leading cause of cancer related mortality in the United Kingdom (UK) in 2012 (1). Evidently, it is a disease of the elderly with men above the age of 75 at higher risk of disease related mortality (2). It has a heterogeneous progression course which varies from slow growing and potentially insignificant to aggressive with serious impact on the patient's health and quality of life. Despite the disease rising prevalence, there has been a significant improvement of its survival rates (3, 4) which can be accounted for the availability of more accurate diagnostic and treatment modalities. With all the promises this may provide to the Pca community, there will be an expected added burden to the limited health care resources with an expected call for effective cost reducing models. One domain of interest is the follow up service provision of patients with stable disease following radical treatment or on the watchful waiting regimen. The current National institute of Health care Excellence (NICE) recommendation is to provide this service in the primary care where appropriate (5). This has not been widely accepted with a palpable concern among primary care physicians and hospital specialists for the lack of the required expertise in the community to meet this expectation (6). The National Prostate Cancer Audit (NPCA) 2014 annual report identified five different models in the prostate cancer service provision in this domain: Consultant-led clinic, Cancer Nurse Specialist (CNS)-led clinic, Telephone clinic, Community-based specialist follow up and Radiographer-led clinic (only for radiotherapy pathway) (7). The community based model had an insignificant contribution in this audit which would support the before mentioned concern about lacking the required experience in the stable prostate cancer domain.

The availability of the experts for problem solving is a common demand in various industries which can be rectified by expert simulating systems i.e. Expert Systems (ES). These are computerised advisory programmes that attempt to imitate or substitute the reasoning process and knowledge of experts in solving specific types of problems (8). They are developed by applying their machine learning techniques where they acquire their knowledge by identifying and following the patterns linking different variable to the outcomes. Alternatively, they can be programmed by the knowledge elicited from the domain experts where a human catalyst is required to digest this knowledge then present it to the system programmer. In fact this is considered to be the bottle neck of the entire system development and its failure will arrest the whole progress (ref). The following stages in the system life cycle would include validating (are we building the right system) and verifying the system (are we building the system correctly). There would be some overlap between both these quality assurance procedures but their main objectives and models are fundamentally different. Despite being a crucial step in the system lifecycle the review of their models in industry demonstrated a lack of consistency and formality (9). In health care, this is even more prominent where the whole system development cycle suffers from live testing shortcoming and lack of adequate system quality assurance (amen). This has led to their lack of uptake in health care with calls for developing standard validation tools satisfying their strict registration criteria with the regulatory bodies such as Medicine and Healthcare products Regulatory Agency (MHRA).

In the above statements we have identified two parallel gaps in health care which we are proposing to cross in this study. First is the lack of decision support tools for simulating Urologist decision making in the community. The second challenge is the development and validation of an ES in the stable prostate cancer domain.

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

TRIAL / STUDY OBJECTIVES AND PURPOSE

PURPOSE

To verify and validate an ES simulating Urology experts in the domain of stable prostate cancer (SPC). The valid system will support community health care workers in providing follow up service led by this study 's ES in SPC.

PRIMARY OBJECTIVE

Test the null hypothesis that the agreement between the ES and the human domain experts is slight to fair (k_0= 0.4) against the alternative that it is better than moderate (k_1 > 0.61).

SECONDARY OBJECTIVES

To verify that the ES is 100 percent reliable i.e. the same clinical input will always trigger the system to produce the exact outputs.

EXPERT SYSTEM STRUCTURE

The main objective of these systems is supporting the user in solving a specific problem, which would usually require an expert judgement. This delicate reasoning process requires storing and manipulating the domain specific knowledge following a structured inferring scheme. For effective communication with the user and the domain expert, the system should be equipped with one or more interfaces to gather the required information.
Thus the standard design of the ES will include these **four main components (10, 11):**

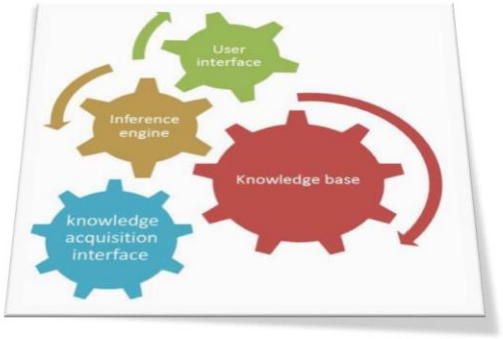


Figure 1: Expert systems component

1. *Knowledge Base:* The core component of the system where all the domain expert knowledge is stored and manipulated for solving the problem. It has a delicate structure that influences the ultimate system model and function. The programmer who is not by default a domain expert codifies the knowledge into an ES framework. An example of this framework is the rule based structure or conditioned phrases (IF-THEN) which we have used to model this study ES.

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

2. *Inference engine:* A separate programme from the knowledge base that contains the system reasoning scheme for solving the domain specified problem. Unlike the Knowledge base, it is more standard in structure and not programmed by any domain specific knowledge.
3. *User Interface:* An application to assist the user in communicating with the system and presenting the specified problem information.
4. *Knowledge acquisition interface:* The system developer interface where knowledge can be updated and corrected in the knowledge base.

DESCRIPTION OF THE PROSTATE CANCER SYSTEM

It is a Clinical Decision Support System (CDSS) which is a web based solution that can be used with current versions of browsers from leading providers such as Microsoft (IE), Google (Chrome), Firefox (Mozilla) and Apple (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. The system uses principally Microsoft technologies including MS SQL Server database, the Microsoft.NET development framework and the C# programming language along with other web technologies like JavaScript and AJAX.

SYSTEM SECURITY AND FUNCTIONAL ANALYSIS

The ES is based on a secure remote server held within the NHS 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 in line with the standards of the Information Governance Statement of Compliance (IG SoC).
The standard system function starts by the authorised user query about a patient from the existing data base. The ES will then access their records on the Hospital Information System (HIS) including their laboratory test results before asking the user to fill a problem specific online form relevant to the present query. The form will only include routine questions enquiring about relevant symptoms as outlined in the NICE 2014 guidelines to support the system accurate decision making. The device has no contact with tissues or body fluids and does not interact directly with the patient. The system suggested outcomes are in the form of an electronic clinic letter suggesting a tailored treatment plan for the individual patient consultation (APPENDIX 1).

TRIAL / STUDY DESIGN

TRIAL / STUDY CONFIGURATION

The SPC ES is already part of an existing CNS 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 existing service is acting as a database and clinic letter generator while the system suggestions being overridden by the CNS. In this study, we will evaluate the system validity by examining the generated electronic letters with the system suggestions after eliminating the CNS suggestions and all patients' identifiers.
These electronic letters are generated for each individual consultation (as per the methods mentioned above in the functional analysis) which will have clear information on the system clinical management and suggested treatment (APPENDIX 1). The CNS will then review and consider these suggestions then produce the final decision by overriding the ES outcome where necessary in the nurse notes section of that letter (Section 5 in the illustration in APPENDIX 1). Thus there is not any compromise of the patients' care in and the system is only acting as a database rather than its intended use as Expert Simulating System.

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

The generated electronic letter (APPENDIX 1) includes five main components:

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

For each patient consultation in this clinic, an electronic letter will be generated and stored on the system data base. The eligible cases will then have their electronic letter anonymised by eliminating all patients' identifiers and the CNS free text entry of their own clinical assessment and plan (Component 5).

The study CM will then examine these cases against the study inclusion criteria and produce a report including a justification for excluded cases from the study (as not enough clinical data, newly diagnosed or treated cases i.e. not SPC etc.).

In order to reach both end points, the study will divide into two branches after the case selections. In the primary end point pathway, the system reliability and precision will be estimated by examining all included cases' variables. Each case electronic letter will have their clinical data inputs (Components 1 and 2) and system output (Components 3 and 4) codified following the coding system outlined in APPENDIX 4. The cases specific codes will be reported in the CASES REPORT FORM (CRF APPENDIX 4) while all cases with similar input (s) are expected to have the same output. Thus all cases with similar inputs will be tabulated and any discrepancies in their outcome will be examined. Any errors detected will be analysed, the system will be corrected then retested by another sample of cases to re-estimate reliability until 100% precision is obtained (primary end point).

In the second part of the study, the included cases will be reviewed by two independent clinical investigators after eliminating the system analysis and suggestions (Components 3 and 4 samples in APPENDIX 3). Each investigator will independently assign an outcome for the presented case according to the disease history, current results and symptoms (Components 1 and 2) blinded from the system recommendation. The CM will document the ratings from the ES original letter and the study investigators in the data report form (APPENDIX 4).

The Kappa statistic will be used to estimate i) the agreement between the investigators and ii) the agreement between each investigator and the ES.

Primary endpoint

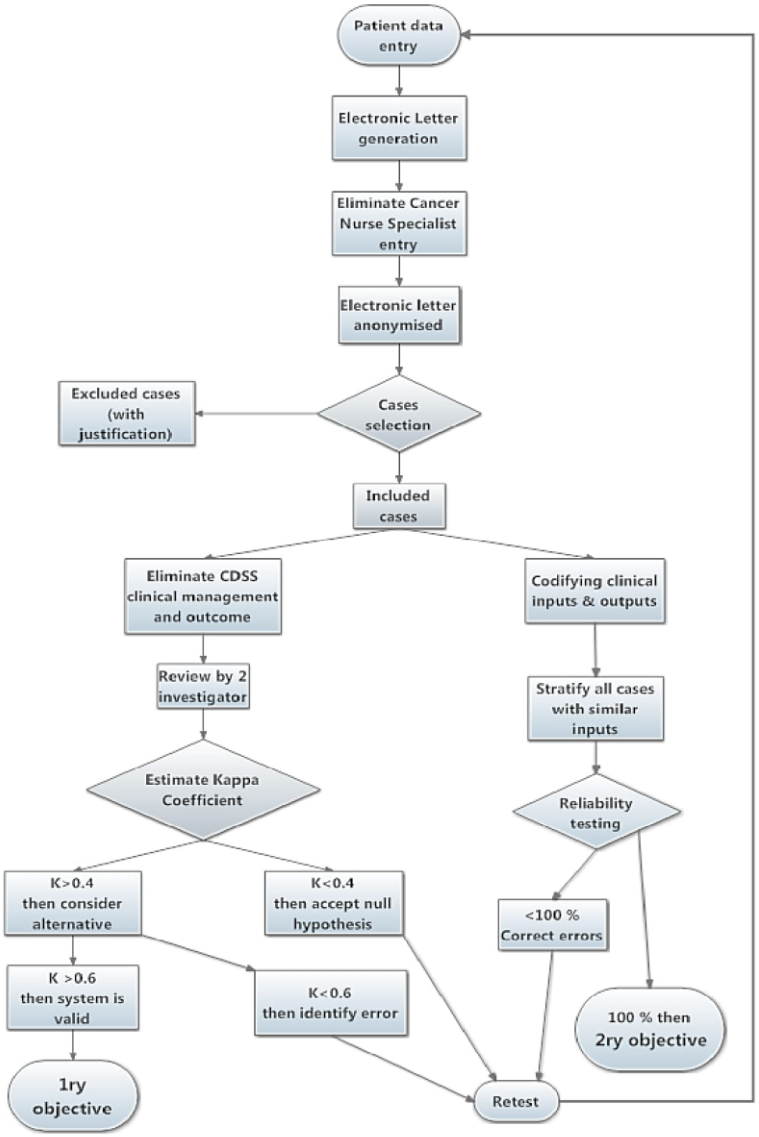
The system is 100% reliable, i.e. similar input must have similar outputs.

Secondary endpoint

Test the null hypothesis that there is agreement between the system and experts is no better than by chance against the alternative that there is adequate agreement between the domain experts and the system to justify its validity ($k > 0.61$).

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

Trial flow chart



This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

Complications from applying the ES in the CNS clinic

1. System errors in any of the components (1 to 4): CNS will have the patient's old records and correct any of these errors in her free text (Component 5) and must report it for tracing this error then correct where appropriate. This should not have any harm on the patient as the CNS is fully qualified in managing stable prostate cancer with direct access to consultant and cancer multidisciplinary team where necessary.
2. Slow system analysis delaying the progress of the clinic: the system has been tested and expected to perform adequately within the NHS settings. In fact in will save the time needed to dictate the clinic letter.
3. Patient not available on the system database, or system missing important information about the patient: this could be due to errors in communication between the HIS and the ES. The CNS will be advised to carry on with her clinic as she would routinely then report this to the Information System engineer.
4. Breach of confidentiality: only authorised personnel can gain access through an encrypted password protected interface. The security is all compliant with the IG SoC standards.

Stopping rules and discontinuation

There are no expected major side effects from applying this system in the CNS clinic to justify discontinuing the study.

BLINDING

The investigators will be blinded from the ES suggestions by omitting Components 3 and 4 of the anonymised electronic letter. Further blinding is deemed unnecessary since it will be include only objective data analysis.

TRIAL/STUDY MANAGEMENT

The Chief Investigator has overall responsibility for the study and shall oversee all study management.

DURATION OF THE TRIAL

The study will run until we identify 200 cases meeting our inclusion criteria which are predicted to take 3 months.

End of the Trial

The trial will end when we reach the primary and secondary end points of our statistical analysis.

SELECTION OF CASES

Eligibility criteria

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

Inclusion criteria

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

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

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

With or without (optional requirement):

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

Exclusion criteria

- Cases with newly diagnosed prostate cancer that are waiting for treatment decision.
- Cases with known prostate cancer that are being evaluated for known recent disease progression (i.e. not SPC).
- Cases with no clinical details of their initial cancer presentation such as Gleason score, PSA, and tumour grade which would influence the follow up decision making.

Informed consent

This study will be collecting prospective data from patients in an existing clinic with no direct patient contact. The data collected are part of the routine clinic as outlined in the NICE recommendations which be further anonymised for the study analysis. Thus patients' consent is deemed unnecessary.

STATISTICS

Methods

Statistical analysis will be carried out by the CM (Mr Hesham Salem) under the supervision of a senior statistician (Dr Lynda Cochrane), using SPSS v22. The primary measure of agreement will be the Kappa statistic (11).

Kappa will be estimated: i) between each expert in the human domain and the ES and ii) between two (or three) experts in the human domain.

Sample size and justification

The sample size required to test the null hypothesis that the agreement between the ES and the human domain experts is slight to fair (k_0= 0.4) against the alternative that it is better than moderate (k_1 > 0.61) will be determined when the frequencies of outcomes in each category is known.

ETHICAL AND REGULATORY ASPECTS

ETHICS COMMITTEE AND REGULATORY APPROVALS

The trial will not be initiated before the protocol has received approval from the Research Ethics Committee (REC), and the respective National Health Service (NHS) Research & Development (R&D) department. Should a protocol amendment be made that requires REC approval, the changes in the protocol will not be instituted until the amendment have been reviewed and received approval

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

from the REC and R&D departments. Minor protocol amendments only for logistical or administrative changes may be implemented immediately; and the REC will be informed.

The trial will be conducted in accordance with the ethical principles that have their origin in the Declaration of Helsinki, 1996; the principles of Good Clinical Practice, and the Department of Health Research Governance Framework for Health and Social care, 2005.

RECORDS

Case Report Forms

The selected cases from the database will be anonymised by the authorised medical team and given a unique identification number for future data auditing if required. The CM will codify the anonymised electronic letters for the included cases (as outlined in the study configuration section) and report these codes in the CRF. All reported data is anonymised and will be used for estimating the ES reliability.

Data Report Form

The clinical monitor will report the ES and clinical investigators' rating in this form for estimating the kappa coefficient for agreement. All the data is also anonymised.

Source documents

Source documents shall be filed at the investigator's site and may include but are not limited to, laboratory results and records. A CRF may also completely serve as its own source data. Only trial staff as listed on the Delegation Log shall have access to trial documentation other than the regulatory requirements listed below.

Direct access to source data / documents

The CRF and all source documents shall be available at all times for review by the Chief Investigator, Sponsor's designee and inspection by relevant regulatory authorities

DATA PROTECTION

All trial staff and investigators will endeavour to protect the rights of the trial's patients' right to privacy and will adhere to the Data Protection Act, 1998. The data will be held securely, in a locked room, or locked cupboard or cabinet. Access to the information will be limited to the trial staff and investigators and relevant regulatory authorities. Computer held data including the trial database will be held securely and password protected. All data will be stored on a secure dedicated web server. Access will be restricted by user identifiers and passwords (encrypted using a one way encryption method).
Information about the trial in the participant's medical records will be treated confidentially in the same way as all other confidential medical information.

QUALITY ASSURANCE & AUDIT

INSURANCE AND INDEMNITY

The University of Nottingham as research Sponsor indemnifies its staff, research participants and research protocols with both public liability insurance and clinical trials insurance. These policies include provision for indemnity in the event of a successful litigious claim for proven non-negligent harm.

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

TRIAL CONDUCT

The clinical monitor shall carry out a site systems audit at least yearly and an audit report shall be made to the Trial Steering Committee.

TRIAL DATA

The Trial clinical monitor shall carry out monitoring of trial data as an ongoing activity. Trial data and evidence of monitoring and systems audits will be made available for inspection by REC as required.

RECORD RETENTION AND ARCHIVING

In compliance with the ICH/GCP guidelines, regulations and in accordance with the University of Nottingham Research Code of Conduct and Research Ethics, the Chief or local Principal Investigator will maintain all records and documents regarding the conduct of the study. These will be retained for at least 7 years or for longer if required. If the responsible investigator is no longer able to maintain the study records, a second person will be nominated to take over this responsibility.

The Trial Master File and trial documents held by the Chief Investigator on behalf of the Sponsor shall be finally archived at secure archive facilities at the University of Nottingham. This archive shall include all trial databases and associated meta-data encryption codes.

DISCONTINUATION OF THE TRIAL BY THE SPONSOR

The Sponsor reserves the right to discontinue this trial at any time for failure to meet expected enrolment goals, for safety or any other administrative reasons. The Sponsor shall take advice from the Trial Steering Committee and Data Monitoring Committee as appropriate in making this decision.

STATEMENT OF CONFIDENTIALITY

Data generated as a result of this trial will be available for inspection on request by the University of Nottingham representatives, the REC, local R&D Departments and the regulatory authorities.

PUBLICATION AND DISSEMINATION POLICY

This study will be a core chapter of Mr Salem's PhD thesis and it will be submitted for publication in peer reviewed scientific journals and meetings.

STUDY FINANCES

Funding source

This study is funded by the urology Department of the Derby Hospitals NHS Foundation Trust.

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

SIGNATURE PAGES

Signatories to Protocol:

Chief investigator: Mr Jonathan Lund

Signature:

Date:

Co-investigator: Dr Daniele Soria

Signature:

Date:

Co-investigator and monitor: Mr Hesham Salem

Signature:

Date:

Clinical investigator: Mr Giacommo Caddeo

Signature:

Date:

Clinical investigator: Mr Jon McFarlane

Signature:

Date:

Statistical advice Dr Lynda Cochrane

Signature:

Date:

Database manager: Mr Kunjan Patel

Signature:

Date:

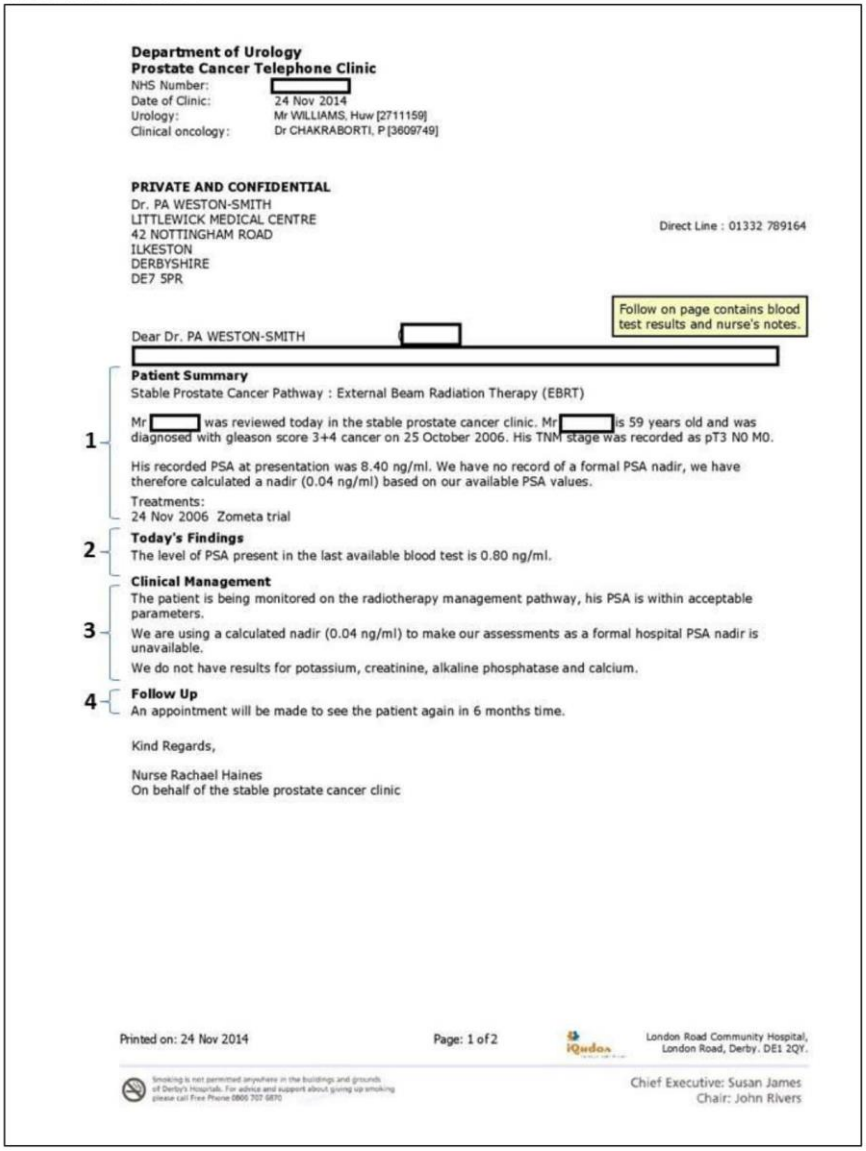
This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

REFERENCES

1. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh JWW, Comber H, et al. Cancer incidence and mortality patterns in Europe: Estimates for 40 countries in 2012. European Journal of Cancer. 2013 4//;49(6):1374-403.
2. Droz J-P, Balducci L, Bolla M, Emberton M, Fitzpatrick JM, Joniau S, et al. Management of prostate cancer in older men: recommendations of a working group of the International Society of Geriatric Oncology. Bju International. 2010 01/12/accepted;106(4):462-9. PubMed PMID: PMC3258484.
3. CancerResearchUK. Statistics and outlook for prostate cancer. http://www.cancerresearchuk.org/about-cancer/type/prostate-cancer/treatment/statistics-and-outlook-for-prostate-cancer: Cancer Research UK Web site; 2014.
4. SEER. SEER stat fact sheets: prostate cancer. National Cancer Institute Web site: National Cancer Institute; 2014.
5. NICE. Prostate cancer: diagnosis and treatment CG175. National Institute for Health and Care Excellence 2014.
6. Watson EK, O'Brien R, Campbell C, Weller D, Neal RD, Wilkinson C, et al. Views of health professionals on the role of primary care in the follow-up of men with prostate cancer. Family Practice. 2011 Dec;28(6):647-54. PubMed PMID: WOS:000297241800009.
7. NCPA. NCPA 2014 Annual Report. http://www.npca.org.uk/reports/: National Prostate Cancer Audit; 2014.
8. Turban E. Decision Support and Expert Systems: Management Support Systems: Prentice Hall PTR; 1990. 864 p.
9. Ammenwerth E, Nykanen P, Rigby M, de Keizer N. Clinical decision support systems: Need for evidence, need for evaluation. Artificial Intelligence in Medicine. 2013 Sep;59(1):1-3. PubMed PMID: WOS:000325194700001.
10. Liao SH. Expert system methodologies and applications - a decade review from 1995 to 2004. Expert Systems with Applications. 2005 Jan;28(1):93-103. PubMed PMID: WOS:000225261500009. English.
11. Lopes M, Marin HD, Ortega NRS. The use of expert systems on the differential diagnosis of urinary incontinence. Revista Da Escola De Enfermagem Da Usp. 2009 Sep;43(3):704-10. PubMed PMID: WOS:000270229400029. English.

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

APPENDIX ONE



Department of Urology
Prostate Cancer Telephone Clinic
NHS Number: [redacted]
Date of Clinic: 24 Nov 2014
Urology: Mr WILLIAMS, Huw [2711159]
Clinical oncology: Dr CHAKRABORTI, P [3609749]

PRIVATE AND CONFIDENTIAL
Dr. PA WESTON-SMITH
LITTLEWICK MEDICAL CENTRE
42 NOTTINGHAM ROAD
ILKESTON
DERBYSHIRE
DE7 5PR

Direct Line : 01332 789164

Follow on page contains blood test results and nurse's notes.

Dear Dr. PA WESTON-SMITH [redacted]

[redacted]

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

1 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.

Treatments:
24 Nov 2006 Zometa trial

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

Clinical Management
3 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 Rachael Haines
On behalf of the stable prostate cancer clinic

Printed on: 24 Nov 2014 Page: 1 of 2 London Road Community Hospital, London Road, Derby. DE1 2QY.

Smoking is not permitted anywhere in the buildings and grounds of Derby's Hospitals. For advice and support about giving up smoking please call Free Phone 0800 707 6870

Chief Executive: Susan James
Chair: John Rivers

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

[redacted]

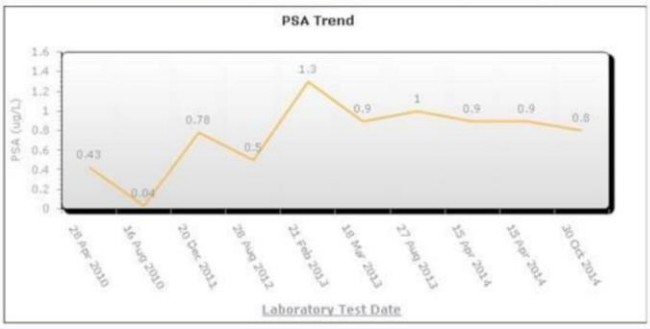
Nurse's Notes:

5 Mr [redacted] is aware of his recent psa. He has no bothersome LUTS, has recently lost weight but this is planned weight loss. He is happy to remain on the 6 monthly telephone clinic.

Recent appointments with associated blood test results:

Blood Test Name	Units	3/5/2013	4/26/2014	11/24/2014
Prostate Specific Antigen (PSA)	ug/L	1.3 (21 Feb 2013)	0.9 (15 Apr 2014)	0.8 (30 Oct 2014)

Key: [-] denotes result not recorded (Date) in parenthesis denotes the lab test date



Printed on: 24 Nov 2014 Page: 2 of 2 London Road Community Hospital, London Road, Derby. DE1 2QY.

Smoking is not permitted anywhere in the buildings and grounds of Derby's Hospitals. For advice and support about giving up smoking please call Free Phone 0800 707 6870

Chief Executive: Susan James
Chair: John Rivers

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

APPENDIX TWO

CASE REPORT FORM (CRF)

Name of the investigator:
Place of the investigation: UoN
Study Identification No.:
Study Duration:
CRF Identification number:

CASE NUMBER	INPUT CODES	OUTCOME CODES	COMMENT

SPC knowledge validation CRF v 1.0 1

CLINICAL MONITOR	
SIGN:	Date ☐☐ / ☐☐☐ / ☐☐
PRINT NAME: HISHAM SALEM	

This form is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

APPENDIX THREE

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

Derby Hospitals NHS
NHS Foundation Trust

Royal Derby Hospital
Uttoxeter Road
Derby
DE22 3NE

Tel: 01332 340131
Minicom: 01332 254944
contactus@derbyhospitals.nhs.uk
www.derbyhospitals.nhs.uk

Re: Mr Study case DB268

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

Mr DB268 was reviewed 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 pT2c N0 M0.

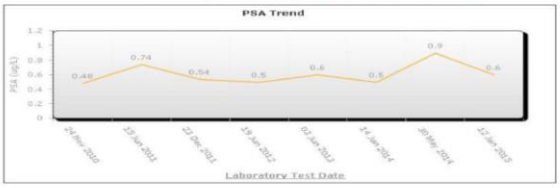
His recorded PSA at presentation was 7.90 ng/ml. His PSA nadir is [0.08 ng/ml].

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	

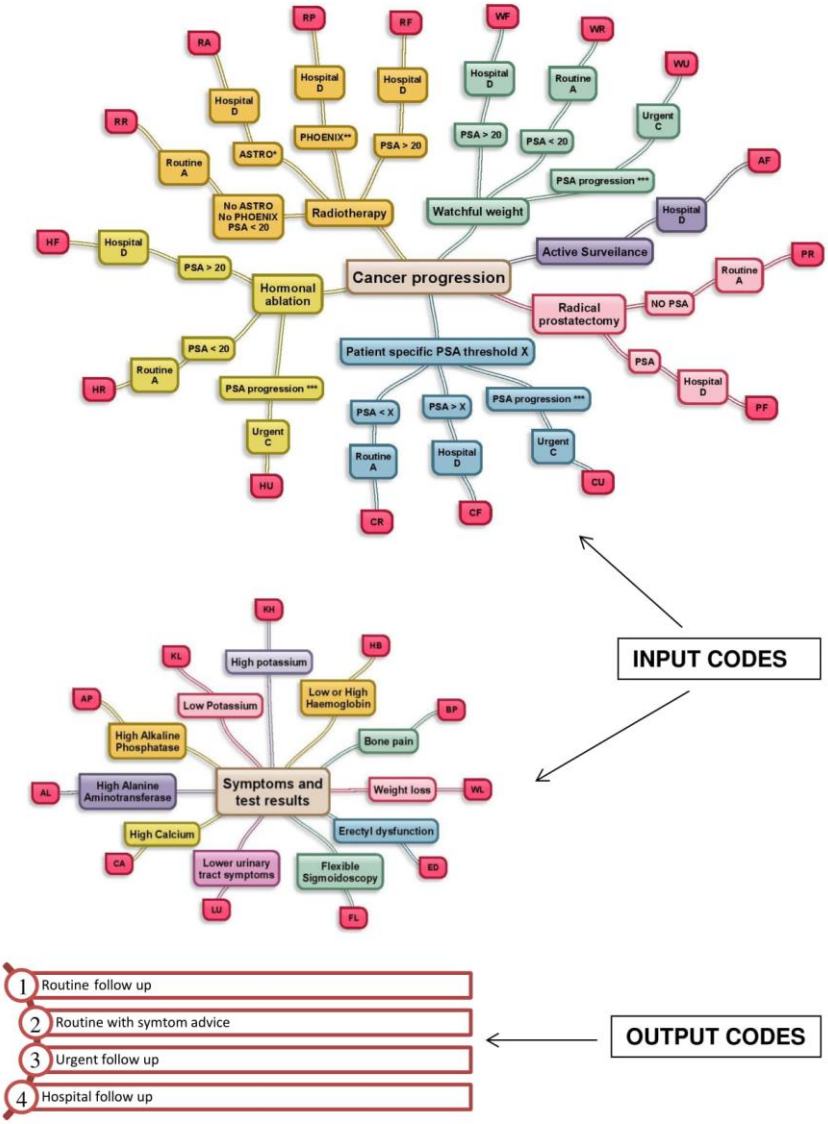
1 = ROUTINE, 2 = ROUTINE WITH ADVICE, 3 = URGENT, 4 = REFER TO HOSPITAL.

NAME	INVESTIGATOR
DATE	
SIGN	

1 | Page

This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham

APPENDIX FOUR



This protocol is confidential and the property of the University of Nottingham. No part of it may be transmitted, reproduced, published, or used by others persons without prior written authorisation from the University of Nottingham