

EAU-EANM-ESTRO-ESUR-SIOG Prostate Cancer Guideline Panel Consensus statements for deferred treatment with curative intent for localised prostate cancer from an international collaborative study (DETECTIVE Study)

Thomas B.L. Lam^{1,2*}, Steven MacLennan^{1*}, Peter-Paul M. Willemse³, Malcolm D. Mason⁴, Karin Plass⁵, Robert Shepherd⁵, Ruud Baanders⁶, Chris H. Bangma⁷, Anders Bjartell⁸, Alberto Bossi⁹, Erik Briers¹⁰, Alberto Briganti¹¹, Karel T. Buddingh¹², James W.F. Catto^{13,14}, Maurizio Colecchia¹⁵, Brett W. Cox¹⁶, Marcus G. Cumberbatch¹³, Jeff Davies¹⁷, Niall F. Davis^{18,19}, Maria De Santis²⁰, Paolo Dell'Oglio^{21,22}, André Deschamps²³, James F. Donaldson^{1,2}, Shin Egawa²⁴, Christian D. Fankhauser²⁵, Stefano Fanti²⁶, Nicola Fossati²⁷, Giorgio Gandaglia²¹, Silke Gillesen^{27,28}, Nikolaos Grivas²⁹, Tobias Gross³⁰, Jeremy P. Grummet³¹, Ann M. Henry³², Alexandre Ingels³³, Jacques Irani³⁴, Michael Lardas³⁵, Matthew Liew³⁶, Daniel W. Lin^{37,38}, Lisa Moris^{39,40}, Muhammad Imran Omar¹, Karl H. Pang¹³, Catherine C. Paterson^{1,41,42}, Raphaële Renard-Penna⁴³, Maria J. Ribal⁴⁴, Monique J. Roobol⁷, Morgan Rouprêt⁴⁵, Olivier Rouvière^{46,47}, Gemma Sancho Pardo⁴⁸, Jonathan Richenberg⁴⁹, Ivo G. Schoots⁵⁰, J.P. Michiel Sedelaar⁵¹, Phillip Stricker^{52,53}, Derya Tilki^{54,55}, Susanne Vahr Lauridsen⁵⁶, Roderick C.N. van den Bergh⁵⁷, Thomas Van den Broeck³⁹, Theodorus H. van der Kwast⁵⁸, Henk G. van der Poel²⁹, Geert J.L.H. van Leenders⁵⁸, Murali Varma⁵⁹, Philippe D. Violette⁶⁰, Christopher J.D. Wallis^{61,62}, Thomas Wiegel⁶³, Karen Wilkinson⁶⁴, Fabio Zattoni⁶⁵, James M.O. N'Dow^{1,2}, Hendrik Van Poppel³⁹, Philip Cornford⁶⁶, and Nicolas Mottet⁶⁷.

** Joint first authors*

Affiliations:

¹ Academic Urology Unit, University of Aberdeen, Aberdeen, UK;

² Department of Urology, Aberdeen Royal Infirmary, Aberdeen, UK;

³ Department of Urology, University Utrecht, Utrecht, The Netherlands;

⁴ Division of Cancer & Genetics, School of Medicine Cardiff University, Velindre Cancer Centre, Cardiff, UK;

⁵ EAU Guidelines Office, Arnhem, The Netherlands;

⁶ Oxford, UK;

⁷ Department of Urology, Erasmus University Medical Centre, Rotterdam, The Netherlands;

⁸ Department of Urology, Skåne University Hospital Malmö, Lund University, Sweden;

⁹ Department of Radiation Oncology, Gustave Roussy Institute, Villejuif, France;

¹⁰ Hasselt, Belgium;

¹¹ Department of Urology, Scientific Institute and University Vita-Salute San Raffaele Hospital, Milan, Italy;

¹² HagaZiekenhuis, The Hague, The Netherlands;

¹³ Academic Urology Unit, University of Sheffield, Sheffield, UK;

¹⁴ Department of Urology, Royal Hallamshire Hospital, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK;

¹⁵ Uropathology Unit, Department of Pathology, Fondazione Irccs Istituto Nazionale dei Tumori di Milano, Milan, Italy;

¹⁶ Department of Radiation Medicine, Zucker School of Medicine, Hempstead, New York;

¹⁷ Wales, UK;

¹⁸ Department of Urology, Beaumont and Connolly Hospitals, Dublin, Ireland;

¹⁹ Royal College of Surgeons in Ireland, Dublin, Ireland;

²⁰ Department of Urology, Charité University Hospital, Berlin, Germany;

- ²¹ Unit of Urology, Division of Oncology, Urological Research Institute; IRCCS Ospedale San Raffaele, Milan, Italy;
- ²² ORSI Academy, Melle, Belgium;
- ²³ Chairman Europa Uomo, Antwerp, Belgium;
- ²⁴ Director of the Asian School of Urology, UAA, Jikei University School of Medicine, 3-25-8 Nishi-Shimbashi, Minato-ku, 105-8461, Tokyo, Japan;
- ²⁵ Department of Urology, University of Zurich, Switzerland;
- ²⁶⁵ Department of Nuclear Medicine, Policlinico S.Orsola, University of Bologna, Italy;
- ²⁷ Division of Cancer Sciences, University of Manchester and The Christie, Manchester, UK;
- ²⁸ Department of Medical Oncology and Haematology, Cantonal Hospital St. Gallen, University of Bern, Bern, Switzerland;
- ²⁹ Department of Urology Netherlands Cancer Institute, Amsterdam, Netherlands;
- ³⁰ Department of Urology, University of Bern, Bern, Switzerland;
- ³¹ Department of Surgery, Central Clinical School, Monash University, Melbourne, Australia;
- ³² Leeds Cancer Centre, St. James's University Hospital, Leeds, UK;
- ³³ Department of Urology, Henri Mondor Hospital, Créteil, France;
- ³⁴ University Hospital of Bicêtre - Paris Sud- Saclay University, 94270 Le Kremlin Bicêtre, France;
- ³⁵ Department of Reconstructive Urology and Surgical Andrology, Metropolitan General, Athens, Greece;
- ³⁶ Department of Urology, Wrightington, Wigan and Leigh NHS Foundation Trust, Wigan, UK;
- ³⁷ Cancer Prevention Program, Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA;
- ³⁸ Department of Urology, University of Washington, Seattle, WA, USA;
- ³⁹ Department of Urology, University Hospitals Leuven, Leuven, Belgium;
- ⁴⁰ Laboratory of Molecular Endocrinology, KU Leuven, Leuven, Belgium;
- ⁴¹ University of Canberra, School of Nursing, Midwifery and Public Health, Canberra, Australia;
- ⁴² Robert Gordon University, School of Nursing and Midwifery, Aberdeen, UK;
- ⁴³ Sorbonne Université, GRC no 5, ONCOTYPE-URO, AP-HP, Academic Department of Radiology, Hôpital Pitié-Salpêtrière- Hôpital Tenon. 75013, 75020 Paris, France;
- ⁴⁴ Uro-Oncology Unit, Hospital Clinic, University of Barcelona, Barcelona, Spain;
- ⁴⁵ Sorbonne Université, GRC n°5, ONCOTYPE-URO, AP-HP, Urology Department, Hôpital Pitié-Salpêtrière, F-75013 Paris, France;
- ⁴⁶ Hospices Civils de Lyon, Department of Urinary and Vascular Imaging, Hôpital Edouard Herriot, F-69437 Lyon, France;
- ⁴⁷ Université de Lyon, Université Lyon 1, Faculté de Médecine Lyon Est, F-69003 Lyon, France;
- ⁴⁸ Department of Radiation Oncology. Hospital de la Santa Creu i Sant Pau. Barcelona, Spain;
- ⁴⁹ Royal Sussex County Hospital Brighton and Brighton and Sussex Medical School, Brighton, Sussex, UK;
- ⁵⁰ Department of Radiology & Nuclear Medicine, Erasmus University Medical Center, Rotterdam, The Netherlands;
- ⁵¹ Department of Urology, Radboud University Medical Centre, Nijmegen, The Netherlands;
- ⁵² Department of Urology, St Vincents Hospital and Campus, Sydney, Australia;
- ⁵³ Garvan Institute of Research, Sydney, Australia;
- ⁵⁴ Martini-Klinik Prostate Cancer Center, University Hospital Hamburg-Eppendorf, Hamburg, Germany;
- ⁵⁵ Department of Urology, University Hospital Hamburg-Eppendorf, Hamburg, Germany;
- ⁵⁶ Department of Urology, Copenhagen University Hospital, Rigshospitalet, 2100 Copenhagen, Denmark;
- ⁵⁷ Department of Urology, St. Antonius Hospital, Utrecht, The Netherlands;
- ⁵⁸ Department of Pathology, Erasmus MC University Medical Center, Rotterdam, The Netherlands;
- ⁵⁹ Department of Cellular Pathology, University Hospital of Wales, Cardiff, UK;

⁶⁰ Departments of Health Research Methods, Evidence and Impact (HEI) and Surgery, McMaster University, Hamilton, ON, Canada;

⁶¹ Division of Urology, Department of Surgery, University of Toronto, Toronto, ON, Canada;

⁶² Department of Urology, Vanderbilt University Medical Center, Nashville, TN, USA;

⁶³ Department of Radiation Oncology, University Hospital Ulm, Ulm, Germany;

⁶⁴ University College London Hospitals, London, UK;

⁶⁵ Urology Unit, Academic Medical Centre Hospital, Udine, Italy;

⁶⁶ Royal Liverpool and Broadgreen Hospitals NHS Trust, Liverpool, UK;

⁶⁷ Department of Urology, University Hospital, St. Etienne, France.

Corresponding author

Dr. Thomas B. L. Lam

Academic Urology Unit, University of Aberdeen,

2nd Floor, Health Sciences Building,

Foresterhill, Aberdeen AB25 2ZD,

United Kingdom

E-mail: thomas.lam@nhs.net or thomasbllam@abdn.ac.uk

Tel: +44 (0) 1224 438130

Fax: +44 (0) 1224 550726

Key words:

Deferred treatment with curative intent; Active surveillance and monitoring; Localised prostate cancer; Eligibility; Follow-up; Reclassification; Outcome measures; Consensus statements; Delphi survey; Consensus group meeting; Clinical practice guidelines

Glossary of terms and acronyms:

EAU: European Association of Urology; EANM: European Association of Nuclear Medicine; ESTRO: European Society for Radiotherapy and Oncology; ESUR: European Association of Urology Section of Urological Research; SIOG: International Society of Geriatric Oncology

Word count

Abstract: 300 words

Body text: 3,562 words

Abstract

Background: There is uncertainty in deferred active treatment (DAT) programmes, regarding patient selection, follow-up and monitoring, reclassification, and which outcome measures should be prioritised.

Objective: To develop consensus statements for all domains of DAT. Design, setting and participants: A protocol-driven, 3-phase study undertaken by the EAU-EANM-ESTRO-ESUR-SIOG Prostate Cancer Guideline Panel in conjunction with partner organisations, including: (1) A systematic review to describe heterogeneity across all domains; (2) A 2-round Delphi survey involving a large, international panel of stakeholders, including healthcare practitioners (HCPs) and patients; and (3) A consensus group meeting attended by stakeholder group representatives. Robust methods regarding what constituted consensus were strictly followed.

Results and limitations: 109 HCPs and 16 patients completed both survey rounds. Of 129 statements in the survey, consensus was achieved in 66 (51%); the rest of the statements were discussed and voted on in the consensus meeting by 32 HCPs and 3 patients, where consensus was achieved in an additional 27 statements (43%). Overall, 93 statements (72%) achieved consensus in the project. Some

uncertainties remained regarding clinically important thresholds for disease extent on biopsy in low risk disease, and the role of mpMRI in determining disease stage and aggressiveness as a criterion for inclusion and exclusion.

Conclusions: Consensus statements and the findings are expected to guide and inform routine clinical practice and research, until higher levels of evidence emerge through prospective comparative studies and clinical trials.

Patient summary: We undertook a project aimed at standardising elements of practice in active surveillance programmes for early localised prostate cancer because currently there is great variation and uncertainty regarding how best to conduct them. The project involved large numbers of healthcare practitioners and patients using a survey and face-to-face meeting, in order to achieve agreement (i.e. consensus) regarding best practice, which will provide guidance to clinicians and researchers.

1. Introduction

Deferred treatment with curative intent (i.e. deferred active treatment, or DAT) has emerged as a feasible alternative to standard radical interventions for low-risk localised prostate cancer. [1-3] This includes active surveillance or active monitoring, whereby patients are not curatively treated immediately but instead are reassessed and monitored at regular intervals, and involves a choice by a patient following counselling with their physician, and alternative treatment options may be considered at a future timepoint. Large, prospective studies are currently underway and medium-term outcomes appear to be promising. [4, 5] However, clinical practice guidelines (CPGs) [6] often acknowledge the significant heterogeneity inherent in deferred treatment strategies, with protocols differing in patient eligibility, selection and recruitment, disease monitoring and reassessment, outcome definition and measurement, and triggers for reclassification and change in management. In short, there is uncertainty regarding the definition of eligible patients, and the optimum follow-up strategies. Although attempts have been made to standardise definitions and terminology via consensus methods, [7] there have been no successful projects which harness clinical and patient expertise aiming to comprehensively standardise practice.

Consequently, the EAU-EANM-ESTRO-ESUR-SIOG Prostate Cancer Guideline Panel in conjunction with partner organisations (Appendix 1) commissioned and undertook a project to develop consensus statements for DAT. The project was unique and novel in its use of protocol-driven consensus methods. [8] The specific objectives were to achieve consensus on the following domains: (1) Criteria for patient selection, inclusion and exclusion; (2) Nature and timing of investigations and assessments during monitoring and follow-up; (3) Criteria and thresholds for reclassification and change in management; and (4) Type of outcome measures which should be prioritised. The study findings will be incorporated into international CPGs issued by the EAU-EANM-ESTRO-ESUR-SIOG Prostate Cancer Guideline Panel and collaborators, and will guide and inform clinical practice and further research.

2. Material and Methods

The protocol outlining the detailed methods underpinning the project has been published. [8] An overview of the study is depicted in Figure 1. The project was divided into 3 phases, lasting 12 months.

Phase 1 was a systematic review of current DAT practice [9] the results of which are summarized in Table 1 and Supplementary Table 1. The review findings were used to inform a list of statements and organised into domains and sub-domains reflecting aspects of DAT (i.e. patient eligibility and recruitment, follow-up and monitoring, reclassification, and outcome measures).

In Phase 2, the list of statements was incorporated into an online questionnaire as part of a two-round iterative Delphi survey. An international panel of participants including healthcare practitioners (i.e. urologists, medical and clinical/radiation oncologists, radiologists, pathologists, and specialist nurses) and patients were purposefully sampled to participate. The list of organisations which participated is included in Appendix 1. These organisations were targeted owing to the expertise of their membership. Organisations provided participants by either nominating individuals or cascading the invitation to their entire membership. Informed consent was assumed if participants registered and completed the survey.

In the online questionnaire, participants were presented with statements and asked to rate their strength of agreement on a scale of 1 (strongly disagree) to 9 (strongly agree). Participants could also suggest additional statements for incorporation into the following round. In Round 2, participants were provided with information regarding their own score from Round 1 as well as a summary of the scores for the entire cohort, and could either revise or retain their original scores. Thresholds regarding what constituted 'consensus agreement' and 'consensus disagreement' were specified *a priori* [8]. 'Consensus agreement' was defined as $\geq 70\%$ of participants scoring a statement as 'strongly agree' (7-9) **and** $< 15\%$ of participants scoring 'strongly disagree' (1-3). Conversely, 'Consensus disagreement' was defined as statements scored as 'strongly disagree' (score 1-3) by $\geq 70\%$ of participants **and** $< 15\%$ of participants scoring 'strongly agree' (7-9). All other statements not falling in the above categories will be classified as equivocal. The decision to use 70% as a threshold was based on prior studies and consensus methods research. [10-13]

Phase 3 consisted of a one-day, face-to-face consensus group meeting attended by representatives from all stakeholder groups and chaired by a non-voting clinician and non-voting methodologist. Participants were sampled from those who completed both rounds of the Delphi survey. All participants were provided with a personalised print-out containing a reminder of how they scored each statement in both rounds of the Delphi, and were given the summary of group results for all statements. All statements not achieving consensus in Phase 2 were discussed, reviewed and voted upon by participants, using the same consensus thresholds from Phase 2, using live voting software [8]. At the end of Phase 3, a final list of consensus statements organised according to the domains of DAT were ratified by the consensus group participants and project steering group.

3. Results

3.1 Delphi survey

Round 1 of the Delphi survey was generated from the systematic review findings (Appendix 2). 127 statements were organised under the following domains and sub-domains: (1) Patient eligibility, inclusion and exclusion criteria: (a) Age and life expectancy; (b) Risk classification (including D'Amico or EAU risk groups, PSA elements, Gleason sum score/ISUP Grade group, clinical stage, etc.); (c)

Histopathological characteristics (including how biopsy is performed, extent of disease, etc.); and (d) Imaging characteristics (including issues regarding multi-parametric MRI, etc.); (2) Monitoring and follow-up criteria (including issues regarding frequency and nature of PSA testing, repeat biopsy, clinical examination by digital rectal examination, and imaging); (3) Reclassification and change in management criteria and triggers: (a) Patient characteristics; (b) PSA kinetics; (c) Histopathology (including change in grade or disease extent); (d) Clinical examination; (e) Imaging; and (f) Patient preference; and (4) Outcome measures which must be prioritised in DAT programmes (including oncological, functional and quality of life [QoL] outcomes).

A total of 180 healthcare practitioners (HCPs) involved with DAT were identified through international specialist societies (Appendix 1) and invited to participate. 50 patients identified through patient advocacy organisations (Appendix 1) were invited to complete the patient-relevant parts of the survey (i.e. outcome measures which should be prioritised). Two additional statements suggested by participants were added to the questionnaire in Round 2 (Appendix 2), bringing the total number of statements to 129. In total, 126 HCPs (70% of those invited) and 29 patients (58% of those invited) completed Round 1, and 109 HCPs (61% of those invited) and 17 patients (34% of those invited) completed both rounds of the survey. The attrition rates between Rounds 1 and 2 were 14% for HCPs, and 41% for patients. Appendix 3 outlines the list of Delphi participants organised by stakeholder group (i.e. HCPs or patients), and including details such as name, speciality and country of residence for HCPs, and previous treatment, age and country of residence for patients.

Table 2 summarises the characteristics of all Delphi participants completing both rounds of the survey, based on stakeholder groups, speciality (or relevant treatment for patients), age (for patients only) and country of residence. Table 3 summarises the survey results for all statements, organised according to consensus status (i.e. consensus, near consensus, divergent opinions, or equivocal/unclear). In summary, there was consensus on 66 statements (51%) from the Delphi survey. The other remaining 63 statements were brought forward for review, discussion and voting in Phase 3, to see if consensus could be achieved on them.

3.2 Consensus group meeting

The consensus group meeting was held in Amsterdam on 9th November 2018 during the 10th European Multidisciplinary Congress on Urological Cancers (i.e. EMUC 2018). The meeting was attended by 35 voting participants (32 HCPs and 3 patients) and chaired by a non-voting clinician and a non-voting methodologist. Table 4 summarises the characteristics of consensus meeting participants based on stakeholder group, speciality and country of residence. Table 5 summarises the results for all statements reviewed, discussed and voted upon, organised according to consensus status 'yes/no' (i.e. in summary, 27/63 statements (43%) achieved consensus during the meeting).

3.3 Final consensus statements and recommendations from DETECTIVE Study

Table 6 summarises all the consensus statements obtained from all phases of the study. In total, 93 statements out of a total 129 (72%) achieved full consensus. The majority of these were achieved from the Delphi survey (71%), whilst the consensus group meeting contributed 29% to the consensus

statements. 53% of the consensus statements were 'consensus agree' whilst 48% were 'consensus disagree'. Consensus was achieved in at least 65% of statements across all domains across the Delphi and consensus meeting process. Table 7 lists all clinical practice recommendations based on the consensus statements.

4. Discussion

4.1 Principal findings

This project explored and defined key areas of controversy and uncertainty covering all the main domains of deferred active treatment, a large undertaking not previously attempted on this scale using transparent methodology. A mixed methods approach was used to investigate this pressing problem, incorporating a systematic review, a two-round Delphi survey and a face-to-face consensus meeting with international participation from key stakeholders. The systematic review confirmed the scale and scope of the problem, highlighting significant heterogeneity, inconsistency and variability in clinical practice across all domains in contemporary studies of DAT. Given such heterogeneity, it is not surprising to note that currently, there is no conclusive data on how different DAT strategies compare to one another, and which strategy, definition and threshold should be adopted in clinical practice, and in clinical trials. Although several seminal randomised controlled trials investigating the effectiveness of observation [1, 2] or active monitoring [3] as a management strategy for localised prostate cancer in comparison with active curative treatment have been published, these studies do not represent current practice of deferred active treatment, which has continued to evolve over the past 15 years, especially with the introduction of new technology such as mpMRI scan into the patient care pathway, changes in the reporting of prostate cancer grade, and more accurate ways of performing prostate biopsies (including MRI-targeted biopsies or transperineal template biopsies). There is, therefore, an urgent need to provide guidance to clinicians, patients, researchers and policymakers, and in the absence of high levels of evidence, the only available option is to issue consensus statements using robust, transparent and reproducible methods. Our project set out to achieve this objective, and ultimately consensus was achieved in more than 72% of statements covering all the domains of DAT, and the results will provide the basis for international guidance and drive the research agenda for the immediate future. The main recommendations based on the consensus statements are listed in Table 7.

4.2 Relevance and impact of study findings on clinical practice and research

Our study, with participation from healthcare practitioners and patients, has provided the basis for conduct of DAT. Consensus statements represent the lowest level of evidence (i.e. level 5) on the evidence-based medicine hierarchy, [14] but in areas where there is low certainty and conflicting evidence, they represent a pragmatic basis for interim guidance. Consensus statements should be regarded as a starting point for clinicians and researchers to guide studies which will provide higher quality evidence and increase certainty. Evidence is never complete; it is ever-evolving, and correspondingly recommendations require updating as necessary. Using our consensus statements as a basis for informing and guiding the conduct of DAT, there is a need for clinicians to prospectively

collect and audit data on DAT in routine clinical practice, and researchers and trialists to conduct clinical trials or prospective comparative studies so that clinical effectiveness data can be obtained. In this context, initiatives such as PIONEER [15] and the Movember Foundation's Global Action Plan Active Surveillance (GAP3) project which aims to establish a global prospective database [16] represent important initial steps.

Our results may be juxtaposed with those of other studies with overlapping aims. Bruinsma *et al.* [7] used consensus methods to develop statements for active surveillance primarily aimed at standardising terms and definitions. The authors published a list of 61 items as a glossary of terms and definitions, whereas our study provides practical guidance for programmes of DAT. Both studies are complementary. MacLennan *et al.* [12] used similar consensus methods in creating a core outcome set applicable across all interventions, including deferred active treatment. The prioritised outcome measures obtained from our study (i.e. core outcomes for DAT) overlap with MacLennan *et al.*'s core outcome set, providing confidence that men with localised prostate cancer and the healthcare practitioners who treat them, regarded the same outcomes as important in two separate samples. More recently, Merriel *et al.* [17] published consensus statements on current best practice of active surveillance in the UK. The statements were developed by a multi-disciplinary group of 27 members consisting of clinical experts and patient experts, informed by a review of the literature, existing guidelines and protocols used by UK Urology departments, and survey data from men with localized prostate cancer. The final consensus statements were then issued by a subgroup of the panel (n=14) at a face-to-face meeting. There are clear similarities between both projects, with both being informed by a review of the literature, and statements were developed by a multidisciplinary panel of clinicians and patients covering similar domains. However, there are major differences. It was unclear if Merriel *et al.*'s project was based on an *a priori* protocol for the systematic review (e.g. PRISMA) and for the consensus phases; the methods, processes and rules underpinning the consensus process, its definitions and how they were developed and achieved were not described. Our project was more international in scope, involved a larger multidisciplinary panel (n=125) and was protocol-driven. We believe these are essential elements in any consensus endeavour which minimise bias, arbitrariness and subjectivity, whilst enhancing rigour, transparency and reproducibility. Nevertheless, there is overlap between the findings of both projects across all domains, and there are no major contradictory findings; as such both projects could be regarded as complementary.

4.3 Strengths and limitations

The study used robust, transparent and reproducible methods based on an *a priori* protocol. The study was international and contemporary in scope, involving patients and a large panel of healthcare practitioners purposively sampled from a broad range of disciplines, all of whom are stakeholders in DAT. A two-step, multi-phase consensus building process based on an iterative Delphi survey and consensus group meeting using anonymous voting techniques was employed, all of which enhanced internal validity. High external validity was achieved by ensuring that the survey items were informed by a systematic review of the literature, which was undertaken according to PRISMA guidelines. In terms of limitations, the project was designed to be pragmatic and practical for participants. Statements had to be brief and concise, and although participants rated their judgements on a scale, decisions were essentially binary in nature (i.e. disagree or agree). Consequently, it was not possible

to address all elements of uncertainty regarding DAT. In particular, the decision-making process regarding patient inclusion or exclusion or reclassification often involves a complex interplay between multiple factors and variables. The relative weighting placed on each variable as one or more variables change within and across patients, and how this affects the decision-making process for patients and clinicians is difficult to conceptualise and address meaningfully in a consensus-finding study. Secondly, within the healthcare practitioners' group, there was a higher ratio of urologists compared with other specialists, in both the Delphi survey and consensus group meeting. However, this reflects contemporary practice, whereby patients within DAT programmes are managed mostly by urologists. Additionally, there was an unusually high attrition rate within the patient group between Rounds 1 and 2 of the Delphi survey (41%). However, the outcome of all statements rated by patients remained stable between Rounds 1 and 2, hence suggesting that the attrition had minimal impact on the consensus outcome. There is also a small risk of introducing sampling error in terms of failure to achieve a balance between contrasting attitudes regarding active surveillance. However, through purposive sampling of a large number and a wide range of clinical practitioners involved in active surveillance, diverse opinions regarding active surveillance would have been achieved and hence minimising this risk.

The choice of a threshold for defining consensus (i.e. 70% in our study) merits a brief discussion. It may be argued that this is an arbitrary figure. However, our decision to use this threshold was informed by the methodological literature and through experience in previous consensus research conducted by members of the project steering group [12, 13, 18]. Many consensus projects define consensus as $\geq 70\%$ of the participants choosing scores 7-9 and $< 15\%$ choosing scores 1-3 (or vice versa) on a 9-point Likert scale, in order to account for the majority opinion whilst not dismissing divergent opinions [10, 11, 19]. The major emphasis in consensus methodology resources is that any threshold must have been judiciously selected, justified and described *a priori* [20, 21]. A higher threshold of 80% or 90% gives undue influence to outlier opinions and would have significantly reduced the number of items reaching consensus which seriously impairs the study's usefulness in clinical practice and research.

Lastly, the study did not achieve consensus on all statements, with 36 items (28%) failing to reach consensus, although 24 items from this group (i.e. 67% out of the total number of statements not reaching consensus) achieved near-consensus (Table 5). This reflects persisting uncertainty even amongst experts and specialists in the field, which can only be resolved through assessment of robust data from comparative studies from which higher levels of evidence can be obtained.

4.4 Areas for further research

We highlight persisting uncertainty and areas for further study. Firstly, for DAT eligibility, there is a need to improve determination of life expectancy more accurately and on an individualised basis. Presently a combination of approaches and strategies are employed, but they apply on a general rather than an individual level. A potential way forward may include studies exploring the creation of nomograms or actuarial tables integrating essential elements influencing life expectancy, such as age, ethnicity, social class, occupation, family history, specific co-morbidities, smoking status, and so on. Secondly, as our project has shown, certain thresholds remain contentious. For instance, thresholds beyond which disease extent on biopsy ought to lead to exclusion of patients with low-risk disease, or

the role of mpMRI in determining disease stage and aggressiveness as a criterion for inclusion or exclusion into DAT programmes, require data from prospective, well-designed studies, incorporating diagnostic accuracy elements and allowing synthesis of evidence regarding clinical effectiveness. In particular, the definition of 'high disease extent' based on biopsy characteristics remains problematic, although there was consensus on its importance. The role of a negative confirmatory biopsy was also not adequately explored in our study and hence deserves further study. In addition, since decision-making for clinicians and patients regarding DAT should be individualised, there is a need to better understand how the complex interaction between multiple factors influences decision-making, especially in terms of relative weighting placed on different variables and their trade-offs; this could be explored through studies utilising discrete choice experiments. [22] In terms of monitoring and follow-up, there was no consensus regarding the role of per-protocol mpMRI nor per-protocol repeat biopsies (i.e. untriggered), nor on its frequency and timing. The lack of consensus on the need for protocol-mandated (i.e. untriggered) repeat biopsies is particularly striking because many contemporary prospective studies on DAT do include them. Although we found consensus regarding repeat biopsy being required if there was a change in mpMRI, DRE progression or PSA progression, it has to be acknowledged that the sensitivity of these triggers for higher grade disease remains unproven. The evolving role of mpMRI in detecting clinically significant disease in place of biopsy is promising, as are new biomarkers (reviewed in [23]), including serum markers (e.g. Prostate Health Index and 4K score), urinary markers (e.g. Prostate Cancer Antigen 3, or PCA3), and tissue markers (e.g. genomic profiling). Once data on these promising diagnostic interventions mature, future studies should integrate them into nomograms predicting the probability of reclassification. In addition, given the current heterogeneity in practice, there is a need to standardise the risk categories and follow-up strategies in large prospective studies. Lastly, the findings from our study will improve and direct the standardisation of undertaking DAT in routine clinical practice and research. Clinicians should use them to carefully design their DAT protocols such that comparative clinical effectiveness data can be prospectively collected, and the results audited regularly. Researchers should follow our guidance and perform clinical trials or prospective cohort studies comparing different DAT protocols against each other and against immediate curative interventions.

5. Conclusions

The EAU-EANM-ESTRO-ESUR-SIOG Prostate Cancer Guideline Panel, in partnership with other leading guideline authorities and patient advocacy organisations (Appendix 1), undertook an ambitious project using a novel and transparent approach in this setting to develop consensus statements for all domains relating to DAT to standardise clinical practice and research. Protocol-driven, robust and transparent methods were utilised. Consensus was achieved on 93 out of 129 statements (72%), covering the domains of criteria for patient selection, inclusion and exclusion (including patient and disease characteristics, imaging criteria, and type of biopsies), nature and timing of investigations and assessments during period of monitoring and follow-up (including PSA measurements, clinical examination, repeat imaging and repeat biopsies), criteria and thresholds for reclassification and change in management, and type of outcome measures which should be prioritised. The findings will guide and inform routine clinical practice and research by being incorporated into guidelines issued by the EAU-EANM-ESTRO-ESUR-SIOG Prostate Cancer Guideline Panel and partner organisations, until higher levels of evidence emerge through prospective comparative studies and clinical trials.

Collaborators: *The following organisations participated in the DETECTIVE Study:*

EAU – EANM - ESTRO - ESUR - SIOG Prostate Cancer Guidelines Panel
European Association of Urology Research Foundation (EAU RF)
European Urology
EAU Section of Oncological Urology (ESOU)
American Society of Clinical Oncology (ASCO)
American Urological Association (AUA)
European Society for Radiotherapy and Oncology (ESTRO)
European Association of Urology Nurses (EAUN)
Canadian Urological Association (CUA)
International Society of Urological Pathology (ISUP)
Urological Society of Australia and New Zealand (USANZ)
European Society of Urogenital Radiology (ESUR)
Urological Association of Asia (UAA)
American Society for Radiation Oncology (ASTRO)
Europa UOMO
Red Sock Campaign
Movember Foundation

Acknowledgements: The authors are grateful to the following individuals and organisations for the following contributions: R. Bryan Rumble, MSc for reviewing and commenting on the manuscript; American Society of Clinical Oncology (ASCO) for providing members who participated in the survey and consensus group meeting; IPSEN for providing an unrestricted educational grant; however IPSEN did not have any access to the study data nor did they have any control over the final manuscript; European Association of Urology (EAU) for providing an unrestricted educational grant; however the EAU did not have any access to the data nor did they have any control over the final manuscript.

References

1. Albertsen, P.C., J.A. Hanley, and J. Fine, *20-year outcomes following conservative management of clinically localized prostate cancer*. 2005. **293**: p. 2095-2101.
2. Wilt, T.J., et al., *Follow-up of Prostatectomy versus Observation for Early Prostate Cancer*. *N Engl J Med*, 2017. **377**(2): p. 132-142.
3. Hamdy, F.C., et al., *10-Year Outcomes after Monitoring, Surgery, or Radiotherapy for Localized Prostate Cancer*. *N Engl J Med*, 2016. **375**(15): p. 1415-1424.
4. Bul, M., et al., *Active surveillance for low-risk prostate cancer worldwide: the PRIAS Study*. 2013. **63**: p. 597-603.
5. Bokhorst, L.P., et al., *A Decade of Active Surveillance in the PRIAS Study: An Update and Evaluation of the Criteria Used to Recommend a Switch to Active Treatment*. *Eur Urol*, 2016. **70**(6): p. 954-960.
6. Mottet N, B.J., Bolla M, Briers E, Cumberbatch MG, De Santis M, Fossati N, Gross T, Henry AM, Joniau S, Lam TB, Mason MD, Matveev VB, Moldovan PC, van den Bergh RCN, Van den Broeck T, van der Poel HG, van der Kwast TH, Rouvière O, Schoots IG, Wiegel T, Cornford P., *EAU-ESTRO-ESUR-SIOG Guidelines on Prostate Cancer* E.A.o. Urology, Editor. 2018: Arnhem, The Netherlands.

7. Bruinsma, S.M., et al., *Expert consensus document: Semantics in active surveillance for men with localized prostate cancer - results of a modified Delphi consensus procedure*. *Nat Rev Urol*, 2017. **14**(5): p. 312-322.
8. Lam, T.B.L., et al., *Study Protocol for the DETECTIVE Study: An International Collaborative Study To Develop Consensus Statements for Deferred Treatment with Curative Intent for Localised Prostate Cancer*. *Eur Urol*, 2019. **75**(4): p. 699-702.
9. Peter-Paul M. Willemse, T.L., Nicolas Mottet, Cathy Yuan, Karin Plass, James Donaldson, Niall Davis, Paolo Dell'Oglio, Christian Fankhauser, Nikos Grivas, Alexandre Ingels, Michael Lardas, Matthew Liew, Karl Pang, Catherine Paterson, Imran Omar, Fabio Zattoni, Tim Buddingh, *Systematic review of deferred treatment with curative intent for localised prostate cancer to explore heterogeneity of definitions, thresholds and criteria and clinical effectiveness*. 2018, PROSPERO International Register of Systematic Reviews.
10. Avery, K.N.L., et al., *Development of a Core Outcome Set for Clinical Effectiveness Trials in Esophageal Cancer Resection Surgery*. *Ann Surg*, 2018. **267**(4): p. 700-10.
11. Williamson, P.R., et al., *The COMET Handbook: version 1.0*. *Trials*, 2017. **18**(Suppl 3): p. 280.
12. MacLennan, S., et al., *A core outcome set for localised prostate cancer effectiveness trials*. *BJU Int*, 2017.
13. van der Poel, H.G., et al., *Sentinel node biopsy for prostate cancer: report from a consensus panel meeting*. *BJU Int*, 2017. **120**(2): p. 204-211.
14. Jeremy Howick, I.C.J.L.L., Paul Glasziou, Trish Greenhalgh, Carl Heneghan, Alessandro Liberati, Ivan Moschetti, Bob Phillips, Hazel Thornton, Olive Goddard and Mary Hodgkinson. *The Oxford Levels of Evidence 2*. 2016; Available from: <https://www.cebm.net/index.aspx?o=5653>.
15. PIONEER. *PIONEER: The European network of excellence for big data in prostate cancer homepage*. 2019 [cited 2019 20/05/2019]; Available from: <https://prostate-pioneer.eu/>.
16. Bruinsma, S.M., et al., *The Movember Foundation's GAP3 cohort: a profile of the largest global prostate cancer active surveillance database to date*. *BJU Int*, 2018. **121**(5): p. 737-744.
17. Merriel, S.W.D., et al., *Best practice in active surveillance for men with prostate cancer: a Prostate Cancer UK consensus statement*. *BJU Int*, 2019. **124**(1): p. 47-54.
18. MacLennan, S., et al., *A randomized trial comparing three Delphi feedback strategies found no evidence of a difference in a setting with high initial agreement*. *J Clin Epidemiol*, 2018. **93**: p. 1-8.
19. McNair, A.G., et al., *Core Outcomes for Colorectal Cancer Surgery: A Consensus Study*. *PLoS Med*, 2016. **13**(8): p. e1002071.
20. Rowe, G., Wright, G., and McColl, A., *Judgment change during Delphi-like procedures: The role of majority influence, expertise, and confidence*. *Technological Forecasting & Social Change* 2005. **72** p. 377–399.
21. Boulkedid, R., et al., *Using and reporting the Delphi method for selecting healthcare quality indicators: a systematic review*. *PLoS One*, 2011. **6**(6): p. e20476.
22. Sculpher, M., et al., *Patients' preferences for the management of non-metastatic prostate cancer: discrete choice experiment*. *Bmj*, 2004. **328**(7436): p. 382.
23. Loeb, S., et al., *Active surveillance for prostate cancer: a systematic review of clinicopathologic variables and biomarkers for risk stratification*. *Eur Urol*, 2015. **67**(4): p. 619-26.