Author/s:
Evans, SM;Murphy, DG;Davis, ID;Sengupta, S;Borzeshi, EZ;Sampurno, F;Millar, JL

Title:
Interpolation to define clinical tumor stage in prostate cancer using clinical description of digital rectal examination

Date:
2018-10-01

Citation:
Evans, S. M., Murphy, D. G., Davis, I. D., Sengupta, S., Borzeshi, E. Z., Sampurno, F. & Millar, J. L. (2018). Interpolation to define clinical tumor stage in prostate cancer using clinical description of digital rectal examination. Asia Pacific Journal of Clinical Oncology, 14 (5), pp.e412-e419. https://doi.org/10.1111/ajco.12875.

Title page

Title

Understanding characteristics of patients with prostate cancer who do not have a cT documented and whether it can be assigned by other clinicians from DRE notes

Sue M Evans [1]

Declan G Murphy [2, 3]

Ian D Davis [4, 5]

Shomik Sengupta [6]

Ehsan Zare Borzeshi [7]

Fanny Sampurno [1]

Jeremy L Millar [1,8]

Affiliations

1. Department of Epidemiology and Preventive Medicine, Monash University, Australia
2. Division of Cancer Surgery, University of Melbourne, Peter MacCallum Cancer Centre, Victoria, Australia
3. Australian Prostate Cancer Research Centre, Epworth Healthcare, Richmond, Australia
4. Monash University Eastern Health Clinical School, Victoria, Australia
5. Eastern Health, Victoria, Australia
6. Department of Urology, Austin Health, Victoria, Australia
7. Capital Markets Cooperative Research Centre
8. Alfred Health Radiation Oncology, Alfred Health, Victoria, Australia

Corresponding author

A/Professor Sue M Evans
Email: sue.evans@monash.edu
Telephone: +61 (0)439 388 262

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/ajco.12875.

Abstract

Aim: This study aimed to assess characteristics of patients with prostate cancer for whom clinical T stage category (cT) was not documented in the medical record and assess whether specialists had concordant conclusions regarding cT based on digital rectal examination (DRE) notes.

Methods: Data from the Prostate Cancer Outcome Registry-Victoria (PCOR-Vic) were interrogated. Four specialists independently assigned cT to DRE notes. Words, or part thereof, associated with agreement between clinicians were identified.

Results: Of the 10,587 men, cT was documented in 8758 (82.7%) cases. Multivariate analysis indicated that poor cT documentation was associated with older patient age (OR 0.80, 95%CI 0.66-0.99 if 75.1-85 years; OR 0.50, 95%CI 0.36-0.72 if >85 years); diagnosis via transperineal compared to transrectal ultrasound-guided biopsy (OR 0.68, 95%CI 0.51-0.91); diagnosed in a private hospital (OR 0.85, 95%CI 0.75-0.96); and a diagnostic Gleason score of >8 compared to ≤6 (OR 0.59, 95%CI=0.48-0.73). cT was more likely documented in men diagnosed via transurethral resection of prostate (OR 2.06, 95%CI 1.64-2.58) and/or receiving treatment in a radiotherapy centre (OR 3.44, 95%CI 2.80-4.23 for external beam radiotherapy; OR 3.57 95%CI 2.44-5.23 for brachytherapy and OR 1.34, 95%CI 1.06-1.69 for combination surgery and radiotherapy). Agreement in cT assignment ranged from kappa of 0.158 to 0.582. Stem word components in DRE notes associated with poorest level of agreement were "abnorm", "hard", "nodul" and those with highest level of agreement were terms “benign” and “smooth”.

Conclusions: Mode of diagnosis/subsequent treatment, and cancer characteristics impacted cT documentation. Third party interpretation of clinical notes is problematic.

Key words: documentation; DRE; text mining; prostatic neoplasm; clinical stage; inter-observer variability.

Introduction

Good documentation underpins high quality care. Information that is not recorded, or is recorded ambiguously, can result in inefficient, inappropriate or unsafe care. A US study of patients diagnosed with acute coronary syndrome found that high completeness of medical record documentation for patient history and physical examination correlated with improved

use of evidence-based medicine and a 20% reduction in in-hospital mortality rate.[1] Another US study found that poor documentation resulted in 32% of patients undergoing at least one duplicate test, of which 20% were determined to be not clinically indicated.[2]

The need for good documentation is particularly important for men with prostate cancer as they often see multiple specialists before choosing treatment and their care is often discussed by multi-disciplinary teams. Clear documentation of disease stage in a standard format provides a shared understanding among the treating team of the extent of the patient's disease, which is necessary to guide management decisions and optimise care coordination.

The Prostate Cancer Outcomes Registry-Victoria (PCOR-Vic) was established in 2009 as a Clinical Quality Registry.[3] Clinical Quality Registries collect high-quality data that enable clinically-credible quality indicators to be calculated and reported, in order to improve medical care.[4] Pilot work by clinicians leading the PCOR-Vic identified that less than 5% of medical records had the clinical T stage category (cT) recorded.[3]

The cT is used to assist in calculating the risk of disease progression and is generally determined through digital rectal examination (DRE) plus review of radiology results. It is required for many prognostic models.[5, 6] It is particularly important in the calculation of low risk disease using the National Comprehensive Cancer Network (NCCN) prognostic model used by PCOR-Vic, which requires a PSA level, Gleason score and cT to be recorded in order for it to be calculated (Box 1) To combat the low baseline cT documentation rate, "Documentation of clinical TNM (cT) prognostic group in the medical record" was included as a quality indicator. Over three years compliance in recording of cT increased to 82%.[7] However, not having a stage for 18% of men prevented risk adjusted outcome assessment from being undertaken on patients who had a Gleason score of ≤ 6 and a PSA level of <10ng/mL. The impact of poor documentation, aside from the aforementioned quality and safety issues, is that indicators such as "the percent of men with low risk disease are undergoing active surveillance", will likely underestimate both the numerator and the denominator. The aims of this study weres to understand characteristics of men for whom no cT was documented and whether cT could be derived from DRE notes taken by trained data collectors. By understanding characteristics of men for whom no cT was documented, it will assist the registry in understanding the extent to which this would bias findings produced by PCOR-Vic. The rationale for understanding rater agreement for staging based on DRE was to assess whether DRE notes could be used to input missing cT data into the registry.

Methods

Study population

Data for this study were sourced from medical records of men diagnosed with prostate cancer and contributing to the PCOR-Vic, which captures approximately 75% of men diagnosed with prostate cancer in Victoria, Australia.

Data collection

Details of the recruitment strategy used by PCOR-Vic have been previously described.[3] Data collectors collect the cT information and, where it is not explicitly recorded, document the DRE notes verbatim. These DRE notes are provided in a report back to diagnosing clinicians (urologists, radiation oncologists and medical oncologists) with their bi-annual benchmark reports with an invitation to complete the cT and return it to the registry so that the system can be updated. Reports are only generated and sent to clinicians if they have more than 20 men registered in the database. Where clinicians do not return the completed form, the cT is recorded as incomplete.

The PCOR-Vic coordinator selected a sample of cases where cT was not documented from regions in Victoria where the review team did not practice. This approach was taken to avoid any bias associated with reviewers understanding how to interpret the DRE notes from previous contact with the diagnosing clinician.

Data elements

Differences in characteristics between patients where cT was or was not documented were assessed using the following explanatory variables: age in 10-year categories; Gleason score according to the Gleason Grade Group system (grade group 1: ≤6; grade group 2: 3+4; grade group 3:, 4+3; grade group 4: 8; grade group 5: 9-10);[8] PSA level at diagnosis in four categories (<4, 4.01-10, 10.01-20, >20ng/mL), method of diagnosis; type of diagnosing hospital (private/public), location of the diagnosing hospital (metropolitan/regional); and treatment provided in the initial 12-month period. cT documentation was determined to have occurred (1) if the DRE was performed and the data collector recorded the cT in the registry; (2) if the cT was recorded by the diagnosing clinician when DRE notes were returned for a cT to be assigned; or (3) when the medical record clearly stated that no DRE was performed (cTX).

A review team comprising experienced specialists (two urologists, a radiation oncologist and a medical oncologist) was asked to assign a cT to those DRE notes that did not have cT documented by the diagnosing specialist. Each clinician was provided with the list of DRE notes, the NCCN staging criteria without reference to imaging (Table 1) and instructions that (1) If there is any doubt concerning that T, N or M classification to which a particular case should be assigned, then the lower (less advanced) category should be assigned; and (2) if a group could not be assigned, it should be recorded as "Not able to be assigned." No imaging details or patient identifiers were provided to reviewers. Level of agreement among reviewers was determined using two approaches:

1. The first approach converted the prognostic category into the NCCN risk of disease progression group of low, intermediate, high and very high risk ("NCCN risk groups").
2. The second approach classified the prognostic groups as follows: any T1 group; T2 or T2a; T2b; T2c; and T3 group or above. ("Stage risk groups")

Identifying words associated with good and poor agreements

Words making up each DRE note were converted to lower case, after which punctuation, numbers and 'stop' words were removed, and 'stemming' was applied. Stop words are listed on the SMART system (Salton, 1971) and have been previously described by Lewis et al.[9] Stemming removes inflected and derived words to their stem, base or root form. Following this, words with less than three characters e.g. ca, pt and a set of uninformative words ('dre', 'prostat', 'prostatectomi') were removed. The words were then categorised into the following four groups:

1. Reviewers agreed that the cT could not be assigned based on DRE note (AgNoAssign);
2. Reviewers agreed that the DRE note could be assigned (DisNoAssign);
3. Reviewers agreed on the assignment of low, intermediate, and high risk disease (AgAssign);
4. Reviewers disagreed on the assignment of low, intermediate, and high risk disease (DisAssign).

Statistical analysis

Descriptive statistics were used to summarize predictor variables. A chi-squared test was used to analyse the association between predicator variables and cT documentation.

Univariate and multivariate logistic regression modelling was used to assess the impact of predictor variables on cT documentation. The p level was set to <0.05.

A chi-squared test was performed to assess whether the sub cohort of cases reviewed by clinicians was representative of the overall cohort of men for whom no cT was documented. Cohen's Kappa coefficients were then calculated to assess the level of agreement between reviewers for NCCN risk groups and stage risk groups (described above). Reference values described by Altman[10] described the strength of the agreement: <0.20 denotes poor agreement, 0.21-0.40 denotes fair agreement, 0.41-0.60 denotes moderate agreement, 0.61-0.80 denotes good agreement and 0.81-1.00 denotes a very good or almost perfect agreement. The percent of times words were associated with disagreement was calculated as follows:

$$\% = \frac{\boldsymbol{DisNoAssign + DisAssign}}{\boldsymbol{AgNoAssign + DisNoAssign + AgAssign + DisAssign}} \textbf{ x 100}$$

Only words mentioned more than 25 times in the notes were analysed.

This project was reviewed and approved by Monash University Human Research Ethics Committee (CF16/1317 – 2016000697).

Results

There were 10,587 men diagnosed with prostate cancer and included in this study. Of these the cT was recorded as not being assessed (cTx) on 331 (3.1%) men. The cT was recorded clearly in the medical records of 6503 men (61.4%). For a further 1924 men (18.2%) the cT was recorded after clinicians returned the completed report enclosed with their report. For 1829 men (17.3%), the cT remained absent even after a request was made to clinicians to complete details in a report where the DRE notes for these cases were provided to them.

Table 2 provides a summary of the characteristics of the cohort. Almost half the men were aged less than 65 years (47%) and most (70%) recorded a baseline PSA level of less than 10 ng/mL. Sixty-three percent were diagnosed in private hospitals and 71% were diagnosed in a metropolitan hospital. Most (85%) were diagnosed via transrectal ultrasound guided biopsy of the prostate (TRUS), and 41% underwent surgery in the initial 12 months after diagnosis.

Factors associated with documentation of cT

Table 2 provides the characteristics of men for whom cT was and was not recorded, using both univariate and multivariate analysis. In the multivariate analysis, men diagnosed in a

private hospital, via transperineal biopsy (TPB), and those aged over 75 years, with a Gleason score of >8, were less likely to have a cT recorded than those diagnosed in a public hospital, via trans-rectal ultrasound, and aged <65 years, with a Gleason score of ≤6 and undergoing a radical prostatectomy as treatment in the initial 12 months post diagnosis. Men were significantly more likely to have a cT documented if they were diagnosed via TURP, received any radiotherapy in the initial 12 months post diagnosis, or if the diagnosing hospital could not be identified from the medical record; compared to men who were diagnosed via TRUS, had surgery in the initial 12 month period following diagnosis and were diagnosed in a metropolitan hospital.

Agreement in classification of cT from DRE notes

A total of 1271 DRE notes were provided to reviewers for classification. The selected sample (n=1271) was more likely than the complete cohort (n=1829) to include patients without a recorded PSA level (1.1% vs 3.2%, p<0.001), with a Gleason score of >8 (10.3% vs 12.1%, p=0.04), and diagnosed in a private hospital (80.2% vs 70.1%, p=0.02). For all other variables there was no significant difference between the sample group and the entire cohort.

There was agreement on the classification, or inability to classify, in 902/1271 (71.0%) of DRE notes when applying the NCCN prognostic risk model. 6 Of those where there was agreement, 853/902 (94.6%) were in the low risk group, 9/902 (1.0%) were in the intermediate risk group, 4/902 (0.4%) were in the high risk group and 36/902 (4.0%) were unable to be classified.

There was agreement among all reviewers on the risk categories in only 367/1271 (28.9%) of DRE notes when applying the stage risk model. Of those where there was agreement, 199/367 (54.2%) were in the T1 group, 119/367 (32.4%) were in the T2a group, 9/367 (2.4%) were in the T2c group, 4/367 were in the T3 or above group and 36/367 (0.8%) were unable to be classified.

Table 3 describes the level of agreement between reviewers in classifying the NCCN and Stage risk models. Agreement was highly variable, with kappa scores ranging from 0.158 for agreement between a urologist and medical oncologist using the Stage risk model, to 0.601 for agreement between a urologist and radiation oncologist, also using the Stage risk model. Across all cross-comparisons the radiation oncologist was the most concordant with other disciplines.

Words associated with high and low level of agreement

There were 899 unique words in all DRE notes. This was reduced to 399 words following removal of punctuation, numbers, stop words, words of less than three characters and uninformative words and following stemming of words. There were 29 words mentioned 25 times or more in the DRE notes.

Table 4 provides a summary of the words commonly associated with high and low level of agreement in assignment of NCCN risk group by clinician reviewers. In DRE notes containing stem words 'abnorm', 'hard' and 'nodul' there was disagreement in the assignment of the risk group to which the patient should be assigned on more than 50% of occasions. There was a high level of agreement (90% or greater) when DRE notes contained stem words 'benign', 'smooth', 'base' and 'slight'.

Discussion

This study identified that cT was less likely to be recorded in older men (aged >75 years), those diagnosed via TPB, diagnosed with a high Gleason score (>8) and on active surveillance or watchful waiting compared with men aged less than 65 years, diagnosed via TRUS, with low risk disease (Gleason grade <6) and having surgery as first-line treatment. We found that there was only modest agreement between clinicians when classifying cT from narrative DRE notes. Commonly used words in DRE notes associated with a low level of agreement between clinicians include 'hard' 'nodular' and 'lobe.' It is likely that the noun 'lobe' is a confounder, in that it is associated with poor agreement because it is used in the same sentence as adjectives that cause confusion among clinicians. The word most commonly used in DRE notes in this cohort was 'firm'. In 22% of DRE notes containing the word 'firm' there was disagreement as to whether this was associated with low, intermediate or high risk disease or whether cancer was present at all.

A concerted effort was undertaken by the registry when it initially began operation to increase cT documentation from its low baseline of 5% to the level of 83% reported in this study. Miller et al examined medical records of 2604 men from 770 facilities and found that 66% of surgical patients and 93% of radiotherapy patients had cT documented.[11] A quality improvement study undertaken in Michigan demonstrated a baseline cT documentation rate of 58% (range from 19% to 96%) when medical records of 491 men with prostate cancer receiving care across 12 practices were examined.[12] Documentation of cT was assessed in patients with head and neck, rectal and bladder cancers contributing to the National Cancer

Database from 1998 to 2011. Baseline rates were between 62 % for rectal cancer to 90% for head and neck cancer and rates increased across each tumour stream over the 13 years.[13].

The implication of having 29% of men for whom a risk category could not be assigned has operational implications for registry reports, but the clinical implication is less clear. From a database perspective, these men cannot be included in risk-adjusted reports, meaning that the elderly will be under-represented. Conversely, reports from patients who go on to have radiotherapy will be more accurate as cT is significantly more likely to be captured in this population. Our experience is that the clinical information systems developed for use in radiotherapy centres routinely contain a mandatory data field for documentation of cT stage but that this capability is not routinely available in clinical information systems used by urologists in private practice or in hospital information systems. Radiotherapy treatment templates guide documentation and ensure that elements such as cT cannot be ignored.[14]

The clinical implication of poor cT documentation in men who tend to be older (>75 years), have clinically obvious high risk disease (Gleason score >8), and are managed in private practice is uncertain. It may be that their management path is obvious and so not recording cT is inconsequential. However, of some concern are those with 3+4 disease, are diagnosed via TP and are managed with active surveillance. Not definitively assigning a cT to men in this group might impact how well they are monitored. This requires further examination.

Results demonstrating that men undergoing EBRT, EBRT and surgery and brachytherapy were significantly more likely than those receiving surgery as monotherapy to have a cT documented suggests that urologists are less likely to document cT compared with radiation oncologists. This may reflect a difference in training in radiation oncology, where the importance of staging and documentation is emphasised and is a principle of their practice, which focuses entirely on cancer treatment. Urologist training and experience is much more diverse, and for many, fundamentals of basic clinical oncology may not be as deeply embedded in vocational training. Secondly, for radiation oncologists the risk grouping of a patient is deterministic for the treatment options considered. Whether or not radiation oncologists include androgen deprivation therapy with external beam radiation therapy will be dependent on disease stage. In contrast, decision-making in surgery is less dependent on the DRE exam result at presentation. It might also be that urologists place greater importance on the pathological T stage and little value on the cT which, despite being a core element of risk models, has increasingly been shown to offer little independent prognostic information in predicting recurrence of localised disease among patients undergoing prostatectomy.[15] The

finding that cT was poorly documented among men diagnosed via TPB may reflect poor documentation by a specific urologist cohort who have adopted TPBs, or perhaps that urologists believe that TPB has superior diagnostic capability,[16] and see no additional benefit in documenting cT. It is logical that cT might be better documented among men undergoing TURP because the assignment of cT category in these men is determined by the histopathology of the TURP specimen.

Textbooks describing the feel of a prostate gland with locally advanced cancer describe it as hard, nodular and irregular.[17] These words were associated with poor agreement among reviewers, which might reflect confusion with surrounding words in the DRE notes or unawareness of the general rule 4 of the TNM classification system: "If there is any doubt concerning that T, N or M classification to which a particular case should be assigned, then the lower (less advanced) category should be assigned."[18] The only moderate agreement among urologists with similar level of experience indicates the need for less ambiguous documentation and that there might be benefit in developing a nomenclature of words commonly used in DRE notes. Text mining such as that undertaken in this study could help inform the development of this document.

There were a number of limitations to this study. While the standard evaluation for prostate cancer includes DRE, increasingly magnetic resonance imaging (MRI) is also used alongside the DRE to stage disease, as it has been shown to contribute significant incremental value to the DRE for localisation of disease.[19] Urologists may have been more inclined to agree on a cT had these results been available. However, MRI is not available to all patients while DRE notes are recorded for most.

Recently released guidelines for management of localised prostate cancer recommend that DRE not be undertaken for asymptomatic men as a routine addition to PSA testing by general practitioners, but instead that it be restricted to specialists trained in the technique prior to consideration for biopsy.[20] A challenge for the PCOR-Vic is to ensure that specialists with these skills adequately record the results using the TNM classification system so that other specialists can understand their meaning and the registry can report risk-adjusted outcomes.

In the short term, based on the findings from this study, PCOR-Vic will instruct data collectors to assign a stage T1c (low risk) to men for whom the diagnosing clinician has recorded that the tumour is benign and/or smooth. However, in the medium term, improvements in quality of care for prostate cancer require that there should be greater

emphasis on ensuring that clinical information systems force documentation of critical data fields, and that fields such as cT, which play such an important role in development of these performance reports, are not discretionary. Consistency of assessment and use of common descriptive terms are also necessary. The impact of not having a shared understanding between clinicians of the extent of cancer in the prostate is that it may result unnecessary and costly additional tests and, at worse, inappropriate and unsafe care.

Acknowledgement

Data involved in this publication has been obtained from the Prostate Cancer Outcome Registry-Victoria which is funded by the Movember Foundation. SE is supported by a Monash Partners Academic Health Science Centre Clinical Research Fellowship. IDD is supported by a National Health and Medical Research Foundation (NHMRC) Practitioner Fellowship (APP1102604)

References

[1] Dunlay SM, Alexander KP, Melloni C, Kraschnewski JL, Liang L, Gibler WB, et al. Medical records and quality of care in acute coronary syndromes: results from CRUSADE. *Archives of Internal Medicine* 2008;168:1692-8.

[2] Stewart BA, Fernandes S, Rodriguez-Huertas E, Landzberg M. A preliminary look at duplicate testing associated with lack of electronic health record interoperability for transferred patients. *Journal of the American Medical Informatics Association* 2010;17:341-4.

[3] Evans S, Millar J, Wood J, Davis I, Bolton D, Giles G, et al. The Prostate Cancer Registry: monitoring patterns and quality of care for men diagnosed with prostate cancer. *British J Urology International* 2012;111:E158-66.

[4] Emilsson L, Lindahl B, Koster M, Lambe M, Ludvigsson JF. Review of 103 Swedish Healthcare Quality Registries. *Journal of Internal Medicine* 2015;277:94-136.

[5] Lowrance WT, Scardino PT. Predictive models for newly diagnosed prostate cancer patients. *Reviews in Urology* 2009;11:117-26.

[6] National Comprehensive Cancer Network. NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines): Prostate Cancer Version 1.2018

[7] Evans S, Millar J, Frydenberg M, Murphy D, Davis ID, Kearns P, et al. Positive surgical margins: rate, contributing factors and impact on further treatment: findings from the Prostate Cancer Registry. *British J Urology International* 2014;114:680-90.

[8] Shahian DM. Public reporting of cardiac surgery performance: introduction. *Ann Thorac Surg* 2011 Sep;92(3 Suppl):S1.

[9] Lewis DD, Yang Y, Rose T, Li F. RCV1: A New Benchmark Collection for Text Categorization Research *Journal of Machine Learning Research* 2004;5:361-97.

[10] Altman D. Practical Statistics for Medical Research. London1991.

[11] Miller D, Spencer B, Ritchey J, Stewart A, Dunn R, Sandler H, et al. Treatment choice and quality of care for men with localized prostate cancer. *Medical Care* 2007;45:401-9.

[12] Filson CP, Boer B, Curry J, Linsell S, Ye Z, Montie JE, et al. Improvement in clinical TNM staging documentation within a prostate cancer quality improvement collaborative. *Urology* 2014;83:781-6.

[13] Reddy S, Handorf E, Smaldone M, Ridge J, Farma J, Sigurdson E. Trends of clinical staging for cancer patients at primary presentation: Have we improved over time? *J Clin Oncol* 2015;33.

[14] Bowman S. Impact of electronic health record systems on information integrity: quality and safety implications. *Perspectives in Health Information Management* 2013:10:1c.

[15] Reese AC, Cooperberg MR, Carroll PR. Minimal impact of clinical stage on prostate cancer prognosis among contemporary patients with clinically localized disease. *J Urology* 2010;184:114-9.

[16] Mukamel D, Mushlin A. Quality of care information makes a difference: an analysis of market share and price changes after publication of the New York State cardiac surgery mortality rates. Medical Care. 1998;36:945-54.

[17] Murtagh J (ed). General Practice. North Ryde, NSW: McGraw-Hill Australia; 2015.

[18] American Joint Committee on Cancer (AJCC). AJCC Cancer Staging Manual. 8th Edition New York, NY: Springer Science + Business Media; 2016.

[19] Mullerad M, Hricak H, Kuroiwa K, Pucar D, Chen HN, Kattan MW, et al. Comparison of endorectal magnetic resonance imaging, guided prostate biopsy and digital rectal

examination in the preoperative anatomical localization of prostate cancer. *Journal of Urology* 2005:174:2158-63.

[20] Prostate Cancer Foundation of Australia, Cancer Council Australia. PSA Testing and Early Management of Test-detected Prostate Cancer- clinical practice guidelines. An overview of the recommendations approved by the National Health and Medical Research Council. Available at: http://www.prostate.org.au/media/611412/PSA-Testing-Guidelines-Overview.pdf2015.

Box 1: NCCN risk groups for disease progression[6]

Risk for disease progression	Criteria
Low risk	T1-T2a ***AND*** Gleason ≤ 6/ grade group 1 ***AND*** PSA< 10ng/mL
Intermediate risk	T2b-T2c ***OR*** Gleason = 7 /grade group 2 or 3 ***OR*** PSA 10-20 ng/mL
High risk	T3a ***OR*** Gleason 8-10 / grade group 4 ***OR*** PSA >20ng/mL
Very high risk	T3b-T4 ***OR*** Gleason 8-10/ grade group 4 or 5 ***OR*** PSA > 20ng/mL
Regional	Any T, N1, M0
Metastatic	Any T, Any N, M1

Table 1: Staging criteria (without reference to imaging) used to classify prostate cancer disease

Prognostic Group	Primary tumour clinical stages only (T)	NCCN Risk categories	Stage risk categories
Tx	Primary tumour cannot be assessed	Unable to assign	Unable to assign
T1	Clinically unapparent tumour neither palpable nor visible by imaging	Low risk disease	T1 group
T1a	Tumour incidental histological finding in 5% or less of tissue resected		
T1b	Tumour incidental histological finding		
T1c	Tumour identified by needle biopsy		
T2	Tumour confined within prostate		T2a group
T2a	Tumour involved one-half of one lobe or less		
T2b	Tumour involved more than one-half of one lobe but not both lobes	Intermediate risk disease	T2b group
T2c	Tumour involves both lobes		T2c group
T3	Tumour extends through the prostate capsule	High risk disease	≥T3 group
T3a	Extracapsular extension (unilateral or bilateral)		
T3b	Tumour invades seminal vesicle(s)	Locally advanced/very high risk	
T4	Tumour is fixed or invades adjacent structures other than seminal vesicles		

Table 2: Characteristics of men for whom a clinical T (cT) stage was or was not documented, and factors associated with cT documentation. Statistically significant findings are in bold type.

Factors		cT documented	cT not documented	Total	P value	Univariate	Multivariate		
		8758 (82.7)	1829 (17.3)	10,587		Odds ratio	95% CI	Odds Ratio	95% CI
Age					P<0.001				
	<65	4135 (82.7)	867 (17.3)	5002		1.00		1.00	
	65-75	3354 (83.5)	664 (16.5)	4018		1.06	0.95-1.18	0.93	0.81 - 1.04
	75.1-85	1112 (82.7)	230 (17.1)	1342		1.01	0.86-1.18	**0.80**	**0.66 - 0.98**
	>85	157 (69.8)	68 (30.2)	225		**0.48**	**0.36-0.65**	**0.50**	**0.36 - 0.72**
PSA at diagnosis					P=0.003				
	≤ 4	1373 (81.7)	308 (18.3)	1681		1.00		1.00	
	4.01-10	4786 (83.0)	980 (17.0)	5766		1.10	0.95-1.26	1.13	0.97 - 1.31
	10.01-20	1431 (85.1)	250 (14.9)	1681		**1.28**	**1.06-1.54**	1.21	0.99 - 1.48
	>20	941 (80.3)	231 (19.7)	1172		0.91	0.75-1.10	0.89	0.70 -

									1.12
	Not available	227 (79.1)	60 (20.9)	287		0.85	0.62-1.16	0.71	0.50 - 1.00
Gleason grade group					P=0.007				
	1: ≤6	3043 (83.6)	597 (16.4)	3640		1.00		1.00	
	2: 3+4	2625 (81.8)	583 (18.2)	3208		0.88	0.78-1.00	**0.78**	**0.68 - 0.90**
	3: 4+3	1242 (83.7)	242 (16.3)	1484		1.01	0.86-1.19	0.84	0.70 - 1.00
	4: 8	889 (84.4)	164 (15.6)	1053		1.06	0.88-1.28	0.88	0.71 - 1.08
	5: >8	863 (79.3)	225 (20.7)	1088		**0.75**	**0.63-0.89**	**0.59**	**0.48 - 0.73**
	Not available	96 (84.2)	18 (15.8)	114		1.04	0.63-1.74	1.38	0.76 - 2.50
Method of diagnosis					P<0.001				
	TRUS	7440 (83.1)	1516 (16.3)	8956		1.00		1.00	
	TURP	866 (85.4)	148 (14.6)	1014		1.19	0.99-1.43	**2.06**	**1.64 - 2.58**
	TPB	26 1 (69.8)	113 (30.2)	374		**0.47**	**0.37-0.59**	**0.68**	**0.51 - 0.91**

	Other	162 (79.8)	41 (20.2)	203		0.80	0.57-1.14	0.71	0.45 - 1.12
	Not available	29 (72.5)	11 (27.5)	40		0.53	0.27-1.08		
Diagnosing hospital type				P<0.001					
	Public	2755 (85.1)	483 (14.9)	3238		1.00		1.00	
	Private	5418 (80.9)	1283 (19.1)	6701		**0.74**	**0.66-0.83**	**0.85**	**0.75 - 0.96**
	Not stated	585 (90.3)	63 (9.7)	648		**1.63**	**1.23-2.15**	0.33	0.08 - 1.43
Diagnosing hospital location					P<0.001	9885			
	Metropolitan	6115 (81.8)	1357 (18.2)	7472		1.00		1.00	-
	Regional	2006 (83.1)	407 (16.9)	2413		1.09	0.97-1.24	0.91	0.80 - 1.04
	Not available	637 (90.7)	65 (9.3)	702		**2.17**	**1.67-2.83**	**4.74**	**1.14 - 19.74**
Treatment received					P<0.001	9939			
	RP	3464 (80.2)	853 (19.8)	4,317		1.00		1.00	
	AS/WW	1439	395	1834		0.90	0.78-1.03	**0.7**	**0.64**

		(78.5)	(21.5)					**5**	**- 0.88**
	EBRT	**1872 (92.8)**	**145 (7.2)**	2017		**3.18**	**2.64-3.83**	**3.4 4**	**2.80 - 4.23**
	RP+ EBRT	**534 (84.2)**	**100 (15.8)**	634		**1.31**	**1.04-1.65**	**1.3 4**	**1.06 - 1.69**
	Brachy	**476 (94.1)**	**30 (5.9)**	506		**3.91**	**2.68-5.70**	**3.5 7**	**2.44 - 5.23**
	ADT	**414 (73.4)**	**150 (26.6)**	564		**0.68**	**0.56-0.83**	0.9 0	0.69 - 1.16
	Other (Chemo, HIFU)	53 (79.1)	14 (20.9)	67		0.93	0.51-1.69	1.1 0	0.56 - 2.02
	Not available	506 (78.1)	142 (21.9)	648		0.88	0.72-1.07	0.9 7	0.79 - 1.21

Legend: ADT: androgen deprivation therapy; AS Active surveillance; Chemo: chemotherapy; EBRT: external beam radiation therapy; HIFU: high-intensity focused ultrasound; RP: radical prostatectomy; TPB: transperineal biopsy; TRUS: transrectal ultrasound; TURP: transurethral resection of prostate; WW: watchful waiting

Table 3: Agreement between reviewers in assignment of prostate cancer risk group

	Urologist 1	Urologist 2	Medical oncologist	Radiation oncologist
Urologist 1		0.518* 0.529†	0.301* 0.158†	0.473* 0.556†
Urologist 2			0.347* 0.232†	0.582* 0.601†
Medical oncologist				0.374* 0.243†
Radiation oncologist				

*Level of agreement according to NCCN risk groups of low, intermediate and high risk disease

† Level of agreement according to the following 5 grades: <2a, T2a, T2b, T2c, ≥T3

Table 4: Stem words recorded ≥ 25 times in DRE notes and level of agreement with NCCN risk group assignment by reviewers

Stem word	Agreed that the DRE note could not be assigned (AgNoAssign) n	Agreed that the DRE note could be assigned (DisNoAssign) n	Agreed on the overall assignment of risk disease (AgAssign) n	Agreement in assignment of risk group (low, intermediate, high risk) N	Disagreed on the overall assignment of risk disease (DisAssign) n	Percent of times word associated with disagreement % (95% confidence interval)
abnorm	0	4	10	8, 2, 0	26	75.0 (68.5-87.3)
hard	0	6	38	35, 3, 0	95	72.7 (64.5-79.8)
nodul	0	3	40	39, 1, 0	52	57.9 (47.3-68.0)
malign	0	3	26	23, 1, 2	21	48.0 (28.1-56.8)
lobe	0	12	164	160, 3, 1	130	45.0 (37.9-49.5)
side	0	3	77	76, 0, 1	50	40.8 (30.1-47.4)
irregular	0	2	40	39, 1, 0	24	39.4 (24.9-49.1)
left	0	4	103	99, 3, 1	62	39.1 (29.4-44.4)
area	0	1	24	23, 1, 0	14	38.5 (21.2-52.8)
suspici	0	1	27	26, 1, 0	15	37.2 (21.0-50.9)
palpabl	0	3	36	35, 0, 1	15	33.3 (16.5-41.6)

examin	2	2	31	31, 0, 0	12	29.8 (13.9-40.3)
bilater	0	0	19	15, 4, 0	8	29.6 (13.8-50.2)
large	0	2	44	44, 0, 0	16	29.0 (15.6-38.5)
rectal	1	3	26	25, 0, 1	7	27.0 (7.9-35.1)
size	0	0	35	35, 0, 0	10	22.2 (11.2-37.1)
firm	0	12	427	425, 1, 1	106	21.7 (16.2-23.0)
gland	0	1	106	105, 0, 1	26	20.3 (13.2-27.3)
small	0	1	117	116, 1, 0	27	19.3 (12.6-25.9)
feel	0	5	96	94, 0, 2	14	16.5 (6.8-19.6)
felt	0	2	24	24, 0, 0	2	14.3 (0.8-23.5)
nodul	0	2	159	156, 3, 0	23	13.6 (8.1-18.2)
enlarg	0	3	171	171, 0, 0	19	11.4 (6.0-14.9)
moder	0	2	156	156, 0, 0	18	11.4 (6.2-15.7)
apex	0	2	41	40, 1, 0	3	10.9 (1.4-17.9)
slight	0	1	59	59, 0, 1	5	9.2 (2.5-17.0)
base	0	1	73	72, 1, 0	5	7.6 (2.1-14.2)

smooth	0	3	82	82, 0, 0	3	6.8 (0.7-9.6)
benign	1	2	148	148, 0, 0	2	2.6 (0.02-4.6)