SCANDINAVIAN JOURNAL OF UROLOGY
2020, VOL. 54, NO. 6, 463–469
https://doi.org/10.1080/21681805.2020.1806354

ARTICLE

Histopathological re-evaluations of biopsies in prostate cancer: a nationwide observational study

B. W. H. van Santvoort[a], G. J. L. H. van Leenders[b], L. A. Kiemeney[c,d], I. M. van Oort[d], S. E. Wieringa[a], H. Jansen[a], R. W. M. Vernooij[a], C. A. Hulsbergen-van de Kaa[e] and K. K. H. Aben[a,c]

[a]Research & Development, Netherlands Comprehensive Cancer Organisation, Utrecht, the Netherlands; [b]Department of Pathology, Erasmus MC University Medical Center, Rotterdam, the Netherlands; [c]Department for Health Evidence, Radboud University Medical Center, Radboud Institute for Health Sciences, Nijmegen, The Netherlands; [d]Department of Urology, Radboud University Medical Center, Radboud Institute for Molecular Life Sciences, Nijmegen, the Netherlands; [e]Laboratory of Pathology East-Netherlands, Hengelo, the Netherlands

ABSTRACT

Background: Grading prostate biopsies has an important role in determining treatment strategy. Histopathological evaluations suffer from interobserver variability and therefore biopsies may be re-evaluated.

Objective: To provide insight into the extent of, characteristics associated with and clinical implications of prostate biopsy re-evaluations in daily clinical practice.

Methods: Patients diagnosed with prostate cancer (PCa) by biopsy between October 2015 and April 2016 identified through the Netherlands Cancer Registry were included. The proportion of re-evaluations was assessed and characteristics were compared between patients with and without biopsy re-evaluation. Interobserver concordance of ISUP grade and EAU prognostic risk classification was determined by calculating Cohen's kappa.

Results: Biopsy re-evaluation was performed in 172 (3.3%) of 5214 patients. Primary reason for re-evaluation in patients treated with curative intent was referral to another hospital. Most referred patients treated with curative intent ($n = 1856$) had no re-evaluation (93.0%, $n = 1727$). Patients with biopsy re-evaluation were younger and underwent more often prostatectomy compared to patients without re-evaluation. The disagreement rate for ISUP grade was 26.1% and interobserver concordance was substantial (κ-weighted = 0.74). Re-evaluation resulted in 21.1% ($n = 14$) of patients with localised PCa in a different prognostic risk group. More tumours were downgraded (57.1%) than upgraded (42.9%). Interobserver concordance was very good ($\kappa_{weighted} = 0.85$).

Conclusion: Pathology review of prostate biopsies is infrequently requested by clinicians in the Netherlands but in a non-negligible minority of patients with localised PCa the pathology review led to a change in prognostic risk group which might impact their treatment.

ARTICLE HISTORY
Received 10 April 2020
Revised 17 July 2020
Accepted 1 August 2020

KEYWORDS
Prostate cancer; prostate biopsy; Gleason grading; risk-stratification; interobserver concordance

Introduction

Prostate cancer (PCa) is the most common type of cancer among Western men, accounting for approximately a fifth of all newly diagnosed malignancies in males [1]. Systematic eight- to 12-core transrectal ultrasound guided (TRUS) biopsy is currently the standard approach to obtain prostate samples for the histopathological evaluation of PCa, although the use of MRI and MRI-ultrasound-fusion guided biopsy is becoming more frequently applied [2]. One of the most essential features of the histopathological evaluation is the Gleason score of the prostate tumour. Gleason grading is the most commonly used PCa grading scheme with worldwide acceptance but suffers from interobserver variability. After the 2005 International Society of Urological Pathologists (ISUP) modifications to the Gleason grading system [3], studies showed interobserver agreement values for prostate biopsies ranging from fair to substantial [4–6]. Given the commonly imprecise Gleason grading and to ensure accurate diagnosis, prostate biopsies may be histopathologically re-evaluated by a pathologist.

Together with the Tumour, Node and Metastasis (TNM) classification and Prostate Specific Antigen (PSA) level, Gleason grading is used to determine the patient's prognostic risk group and the most appropriate treatment plan [2,7]. Re-evaluations of Gleason scores may lead to a change in risk classification and thereby to another treatment plan [4]. This may affect patients' survival, quality-of-life and health system expenses. Few studies examined the effect of histopathological re-evaluations of prostate biopsies on treatment decision in daily clinical practice [4,8,9] and none of these studies were population-based.

The current population-based study was conducted to provide insight into the extent of, characteristics associated with and diagnostic and clinical implications of prostate biopsy re-evaluations in daily clinical practice.

CONTACT K. K. H. Aben k.aben@iknl.nl Research & Development, Netherlands Comprehensive Cancer Organisation, Postbus 19079, 3511 DT, Utrecht, the Netherlands

© 2020 Acta Chirurgica Scandinavica Society

Patients and methods

Patients diagnosed with PCa were identified through the population-based Netherlands Cancer Registry (NCR) held by the Netherlands Comprehensive Cancer Organisation (IKNL). Since 1989, all newly diagnosed malignancies in the Netherlands have been recorded in the NCR. Notifications of new cancers are mainly obtained from the nationwide network and registry of histology and cytopathology (PALGA). Notifications of tumours without histological confirmation are obtained from the National Registry of Hospital Discharge Diagnosis. Data concerning patient and tumour characteristics and initial treatment are routinely extracted from patients' medical hospital records by data managers of the NCR. For patients diagnosed with PCa between 1 October 2015 and 16 April 2016, additional detailed data on the diagnostic process and subsequent treatment were collected within the framework of the NCR. This additional data collection was part of the ProZIB study (acronym for: Prostaatkanker Zorg in Beeld or PCa Care Visualized). The aim of this study was to provide insight in the current clinical practice of PCa and to evaluate quality of clinical care in the Netherlands.

For the current study, we included all patients from ProZIB who were diagnosed with PCa by prostate needle biopsy. Men whose prostate biopsies were obtained and evaluated in the context of a clinical trial concerning the diagnostic process of PCa were excluded as these do not correspond with clinical practice. Tumours were staged according to the 7th version of the TNM classification of the International Union Against Cancer [10]. Hospitals were grouped into university hospitals, non-university referral hospitals and community hospitals. Primary treatment was categorised into the following groups: radical prostatectomy, external beam radiation therapy (EBRT) and/or brachytherapy, combination of hormonal- and radiotherapy, hormonal therapy only or combined with chemotherapy, no active therapy and other. Gleason score comprises the sum of the primary Gleason pattern and secondary Gleason pattern. Gleason scores were grouped into ISUP (International Society of Urological Pathology) grades following the 2014 classification [11]. Prognostic risk groups were made based on the classification of the European Association of Urology (EAU) [2]. Localised low risk PCa was defined as cT1-2a, Gleason ≤ 6 and PSA $<$ 10 ng/mL; localised intermediate risk as cT2b or Gleason 7 or PSA 10–20 ng/mL; and localised high risk as cT2c, Gleason $\geq$ 8 or PSA $>$ 20 ng/mL. Locally advanced included cT3-4 or cN + and metastatic disease included all patients with cM1.

We determined the proportion of prostate biopsy re-evaluations and performed descriptive analyses of age at diagnosis, type of biopsy (ultrasound or MRI), hospital of diagnosis, primary treatment and tumour characteristics. Normally distributed continuous variables were compared between men with and without a re-evaluation using Student's *t*-test, non-normally distributed continuous variables using Wilcoxon signed rank test, categorical variables using Pearson's chi-square test or Fisher's exact test if the expected counts for categories were less than five. Interobserver concordance of the ISUP grade was assessed in all patients with a biopsies re-evaluation. Interobserver concordance concerning the EAU risk classification was determined in patients with localised PCa since re-evaluation-based changes in ISUP grade may affect the risk classification in these patients. Cross tables with frequencies were made to visualise concordance. We calculated the unweighted and linear weighted kappa coefficients (κ) with 95% confidence interval (CI). Kappa values of 0.00–0.20 reflect slight agreement; 0.21–0.40 fair agreement; 0.41–0.60 moderate agreement; 0.61–0.80 substantial agreement; and 0.81 and higher reflect very good agreement [12]. Significance of the tests were assessed at the alpha = 0.05 significance level. All statistical analyses were performed using SAS version 9.4 (SAS Institute, Cary, NC).

Results

Between October 2015 and April 2016, 5214 patients were diagnosed with PCa by prostate biopsy. In 3.3% ($n = 172$) a histopathological re-evaluation of the biopsies was performed. The main reason for the re-evaluation in patients treated with curative intent ($n = 139$) was referral to another hospital (92.8%, $n = 129$). However, most of the PCa patients treated with curative intent who were referred to another hospital ($n = 1856$) did not have a prostate biopsies re-evaluation (93.0%, $n = 1727$).

Patients with a biopsy re-evaluation more often underwent radical prostatectomy and less often hormonal therapy compared to patients with no re-evaluation (Table 1). Furthermore, patients whose biopsies were not re-evaluated were older and more often had low risk localised or metastatic PCa. The differences between these groups are likely related to the higher proportion of patients with a re-evaluation referred to another hospital for centralized treatment.

Interobserver concordance of the ISUP grade

There was disagreement on the ISUP grade in 26.1% ($n = 43$) of the patients whose biopsies were re-evaluated (Table 2). Of the discordant patients, 48.8% ($n = 21$) were downgraded, whereas 51.2% ($n = 22$) were upgraded. Substantial interobserver concordance was found with a weighted Cohen's kappa of 0.74 (95% CI = 0.66–0.82).

Higher ISUP grade, mainly group 2, was assigned at the re-evaluation in 17.3% ($n = 9$) of the patients who scored ISUP grade 1 at the initial evaluation. Of the 33 patients with ISUP grade 4 at the initial evaluation, 30.3% ($n = 10$) were downgraded at the re-evaluation, mostly to ISUP grade 3. There were no false positives identified.

Interobserver concordance of the EAU prognostic risk classification

In patients with localised PCa and a re-evaluation, the disagreement rate of the EAU risk classification was 12.1% ($n = 14$, Table 3). The proportion of tumours that was downgraded was slightly higher than the proportion of tumours upgraded (6.9% vs 5.2%). The weighted kappa coefficient was 0.85 (95% CI = 0.77–0.93), which reflects very good

Table 1. Patient, diagnostic, tumour and treatment characteristics of patients with prostate cancer whose biopsies were and were not re-evaluated.

Characteristics[a]	Patients *with* re-evaluation ($n = 172$)	Patients *without* re-evaluation ($n = 5042$)	*p*-value
Patient characteristics, mean (SD)			
Age at diagnosis (years)	66.8 (6.0)	69.3 (7.7)	<0.0001
Diagnostic characteristics, *n* (%)			
Type of biopsy guiding			0.0006
Echo-guided	151 (87.8)	4724 (93.7)	
MRI and MRI-echo-fusion guided	21 (12.2)	279 (5.5)	
Unknown	0 (0.0)	39 (0.8)	
Hospital of diagnosis			<0.0001
University hospital	5 (2.9)	404 (8.0)	
Non-university referral hospital	63 (36.6)	2489 (49.4)	
Community hospital	104 (60.5)	2149 (42.6)	
Tumour characteristics, *n* (%)			
PSA at diagnosis (ng/mL), median (IQR)	8.7 (7.8)	10.8 (20.2)	<0.0001
Missing	3 (1.7)	51 (1.0)	
Clinical TNM stage			0.0008
cT1-2a	83 (48.3)	2325 (46.1)	
cT2b	10 (5.8)	134 (2.7)	
cT2c	30 (17.4)	648 (12.9)	
cT3–4 or cN+	39 (22.7)	1129 (22.4)	
cM+	10 (5.8)	801 (15.9)	
Unknown	0 (0.0)	5 (0.1)	
ISUP grade			<0.0001
1 (3 + 3)	55 (32.0)	1838 (36.5)	
2 (3 + 4)	40 (23.3)	1090 (21.6)	
3 (4 + 3)	29 (16.9)	500 (9.9)	
4 (8)	33 (19.2)	712 (14.1)	
5 (9–10)	10 (5.8)	839 (16.6)	
Unknown	5 (2.9)	63 (1.3)	
EAU prognostic risk group			0.0002
Localised low risk	27 (15.7)	1033 (20.5)	
Localised intermediate risk	38 (22.1)	967 (19.2)	
Localised high risk	53 (30.8)	1039 (20.6)	
Locally advanced	39 (22.7)	1129 (22.4)	
Metastatic	10 (5.8)	801 (15.9)	
Unknown	5 (2.9)	73 (1.5)	
Treatment characteristics, *n* (%)			
Primary treatment			<0.0001
Radical prostatectomy	97 (56.4)	1141 (22.6)	
EBRT and/or brachytherapy	22 (12.8)	659 (13.1)	
Hormonal- and radiotherapy	20 (11.6)	870 (17.3)	
Hormonal therapy only or combined with chemotherapy	10 (5.8)	1061 (21.1)	
No active therapy	19 (11.1)	1257 (24.9)	
Other	4 (2.3)	54 (1.1)	

[a]Results of the first evaluation are displayed.
Abbreviations: SD, standard deviation; IQR, interquartile range; EBRT, external beam radiation therapy.

Table 2. Interobserver concordance of the ISUP grade between the prostate biopsies initial evaluation and re-evaluation.

	Re-evaluation ISUP grade							
Initial evaluation ISUP grade	1 (3 + 3)	2 (3 + 4)	3 (4 + 3)	4 (8)	5 (9 − 10)	Total	Upgrade %	Downgrade %
1 (3 + 3)	43	8	1	0	0	52	17.3	—
2 (3 + 4)	3	32	3	3	0	41	14.6	7.3
3 (4 + 3)	3	3	19	1	3	29	13.8	20.7
4 (8)	1	3	6	20	3	33	9.1	30.3
5 (9 − 10)	1	0	0	1	8	10	—	20.0
Total	51	46	29	25	14	165		
	Frequency missing = 7 (4.1%)							

Shaded boxes indicate cases with identical ISUP grade. Percentage total agreement = 73.9%. Total upgrade = 13.3%, total downgrade = 12.7%. $\kappa_{unweighted}$ = 0.66 (0.57–0.75), $\kappa_{weighted}$ = 0.74 (0.66–0.82).

interobserver concordance. Intermediate risk was assigned at the re-evaluation in 15.4% ($n = 4$) of the patients classified as low risk at the initial evaluation.

Discussion

Based on the current study it can be concluded that a histopathological re-evaluation of the prostate biopsies was performed in a minority of the patients diagnosed with PCa. Interobserver concordance of the ISUP grade was substantial. However, the treatment strategy might change in one out of eight patients with localised PCa as a result of the re-evaluation-based changes in EAU risk classification.

In this study, we tried to identify differences in patients and tumour characteristics between patients with and without a re-evaluation (Table 1). Patients whose biopsies were re-evaluated were younger (66.8 vs 69.3 years), had less often

Table 3. Interobserver concordance of the EAU prognostic risk classification between the prostate biopsies initial evaluation and re-evaluation in patients with localised PCa.

	Re-evaluation risk group					
Initial evaluation risk group	Low	Intermediate	High	Total	Upgrade %	Downgrade %
Low	22	4	0	26	15.4	—
Intermediate	4	31	2	37	5.4	10.8
High	1	3	49	53	—	7.5
Total	27	38	51	116		
	Frequency missing = 7 (4.1%)					

Shaded boxes indicate cases with identical risk group. Percentage total agreement = 87.9%.
Total upgrade = 5.2%, total downgrade = 6.9%. $\kappa_{unweighted} = 0.81$ (0.72–0.90), $\kappa_{weighted} = 0.85$ (0.77–0.93).

metastatic PCa (5.8% vs 15.9%) and underwent prostatectomy more often (65.4% vs 22.6%) compared to patients with no re-evaluation. In the Netherlands, there is an ongoing centralisation of radical prostatectomies. Therefore, patients who underwent surgery were more likely to be referred to another hospital where the surgery was performed. Apparently, biopsies are more often re-evaluated after referral. This is also reflected in the data, as patients with a re-evaluation were more often referred to another hospital for treatment (87.7% vs 39.1%).

Several studies assessed the interobserver reproducibility of the Gleason grading system in prostate biopsies (Table 4). To improve risk stratification, during the 2014 ISUP consensus conference, a categorisation of Gleason scores into fewer and prognostically comparable groups, defined as ISUP grades, was proposed [11]. The study by Ozkan et al. [6] also investigated the interobserver concordance of the ISUP grade in prostate biopsies. They included 34 PCa patients with in total 407 prostate biopsy slides which were all re-evaluated by two pathologists who were blinded to the initial pathology report. Compared to our study, Ozkan et al. reported a higher disagreement rate (48.3% vs 26.1% in our study) and lower interobserver concordance (fair vs substantial in our study). This difference might partly be explained by the heterogeneous study designs. The study of Ozkan et al. was designed in a research setting as a validation study for the interobserver variability, while we aimed to study the implications of re-evaluation in daily clinical practice with a population based series. Consistent with the present results, the study of Berg et al. [4] found disagreement of Gleason groups in 20.4% of the patients ($\kappa = 0.67$). This study was performed in clinical practice and included 350 patients who were all referred to the urology department of a single center. Gleason scores were categorized into groups largely comparable to the ISUP grades, i.e. $\leq$ 5, 6, 7 and 8–10.

Disagreement in our study appeared highest in prostate biopsies with ISUP grade 3 and 4 (Table 2). Almost one third (30.3%) of the patients with ISUP grade 4 at the initial evaluation were downgraded. The most important cause for grading discrepancy might be the interpretation of ill-formed Gleason pattern 4 and cords Gleason pattern 5 structures [13,14]. A three-dimensional microscopic imaging study in fact revealed that both patterns form an architectural continuity, in which setting a cut-off is highly subjective [15]. Previous studies have also found difficulty in distinguishing Gleason score 7 (3 + 4 vs 4 + 3) tumours [4,16,17]. We showed 20.7% downgrading for ISUP grade 3 biopsies compared to only 7.3% downgrading for ISUP grade 2 biopsies. Variability in assignment of Gleason score 3 + 4 vs 4 + 3 may be explained by differences in reporting of either global Gleason score for the entire case or biopsy site, vs worst Gleason score of the biopsy with the highest grade [15]. For instance, in case of two positive biopsies, one with 6 mm Gleason score 3 + 3 = 6 and the other with 2 mm of Gleason score 4 + 3 = 7, pathologists might report and urologists might use for decision-making either the worst grade (4 + 3 = 7) or the global grade (3 + 4 = 7).

Reclassification of the ISUP grade may impact the prognostic risk classification on which treatment decision is based. The shift between ISUP grade 1 and grade 2 or 3 causes a change from low risk to intermediate risk PCa. Low-risk PCa patients may be candidates for active surveillance, whereas patients with intermediate risk PCa are likely to undergo treatment like radical prostatectomy or radiotherapy [2], all of which are associated with side-effects and a reduced quality-of-life. Another clinically important shift is between ISUP grade 2 or 3 and grade 4, as this makes the distinction between intermediate and high risk PCa. Patients with high risk PCa are generally treated more aggressively and have a worse prognosis and lower survival rate [1,18].

Despite the high interobserver concordance of the risk classification, we found that 12.1% of the patients with localised PCa whose biopsies were re-evaluated were either upgraded or downgraded to another risk group. In these patients, treatment strategy potentially changed, which could have affected their prognosis and/or quality-of-life. Besides, re-evaluations may have an impact on the patient-tailored approach without affecting the risk classification. As a result of the re-evaluation, changes in histopathological parameters may influence whether the surgery is performed nerve-sparing or if it is necessary to add hormonal treatment in radiotherapy.

The concordance rate of the EAU risk classification in the current study (87.9%) was higher than reported by most other studies. For instance, the study by Camara-Lopes et al. [19] included 182 patients with PCa who were all referred to the same hospital before undergoing brachytherapy. This study was performed in daily clinical practice and reported a concordance rate of 68.1% ($n = 124$). A study by Thomas et al. [8] observed a management change in 14.8% ($n = 196$) of the patients who were treated in one of four centers with prostate brachytherapy between 1998 and 2005. A fair comparison of the results from these studies with our study is difficult because the other studies were confined to a specific group of patients with PCa who will likely have different cancer characteristics compared to the patient characteristics in

Table 4. The interobserver reproducibility of the Gleason grading system in prostate biopsies, divided by studies conducted before and after the 2005 ISUP modifications to the Gleason grading system.

	Pathologists		Cases	Methods		Results	
Study	No.	Blinded	No.	Comparison standard	Outcome	Exact agreement	Interobserver concordance
Studies conducted *before* 2005 ISUP modifications							
Steinberg et al, 1997 [20]	Unknown	Unknown	390	Against initial evaluation	Gleason groups[1]	62%	—
Lessells et al, 1997 [21]	12	Unknown	46	Against each other	Gleason score Gleason groups[2–5]	—	*Gleason score* Moderate ($\kappa_w = 0.45$) *Gleason groups* Fair/ moderate ($k = 0.37–0.54$)
Wurzer et al, 1998 [22]	8	Unknown	538	Against initial evaluation	Gleason score	60.6%	—
Allsbrook et al, 2001 [23]	41	Unknown	38	Against consensus cases	Gleason groups[1]	—	Moderate ($\kappa = 0.44$)
Allsbrook et al, 2001 [23]	10	Unknown	46	Against each other	Gleason score Gleason groups[1]	*Gleason score* 52%	*Gleason score* Moderate/substantial $\kappa_w = 0.56–0.70$ *Gleason group* Moderate/substantial $\kappa = 0.47–0.64$
Renshaw et al, 2003 [24]	1	Unknown	416	Against initial evaluation	Gleason score Gleason groups[6]	*Gleason score* 59% *Gleason groups* 78.4%	—
Nguyen et al, 2004 [20]	3	Unknown	602	Against initial evaluation	Gleason score Gleason groups[1]	*Gleason score* 56% *Gleason groups 67*%	—
Glaessgen et al, 2004 [25]	4	Unknown	279	Against each other	Gleason score	49%	Moderate ($\kappa_w = 0.51$)
Coard et al, 2004 [26]	1	Yes, to initial evaluation	90	Against initial evaluation	Gleason score Gleason groups[3]	*Gleason score* 60% *Gleason groups* 80%	*Gleason groups* Moderate ($\kappa = 0.54$)
Oyama et al, 2005 [27]	14	Unknown	37	Against consensus cases	Gleason groups[1]	—	*Urological pathologists* Substantial ($\kappa = 0.68$) *General pathologists* Moderate ($k = 0.49$)
Sooriakumaran et al, 2005 [28]	1	Unknown	83	Against initial evaluation	Gleason score	52%	Fair ($k = 0.27$ $k_w = 0.25$)
Melia et al, 2006 [16]	9	Yes, to all other evaluations	81	Against consensus score	Gleason score Gleason groups[1]	*Gleason score* 69% *Gleason groups* 78%	*Gleason score* Slight/ moderate $\kappa = 0.08–0.58$ *Gleason groups* Moderate $\kappa = 0.54$
Griffiths et al, 2006 [29]	24	Unknown	20	Against Consensus score	Gleason score Gleason groups[1]	*Gleason score* 46%	*Gleason groups* Fair ($\kappa = 0.33$)
Thomas et al, 2007 [8]	24	Unknown	1323	Against initial evaluation	Gleason score	74.8%	—
Fine et al, 2008 [30]	Unknown	Unknown	1455	Against initial evaluation	Gleason groups[1]	81.9%	—
Studies conducted *after* 2005 ISUP modifications							
Veloso et al, 2007 [31]	3	Yes, to initial evaluation.	107	Against each other	Gleason score	58.6–69.6%	Fair ($\kappa = 0.36$) (Range $\kappa_w = 0.31–0.44$)
Brimo et al, 2010 [9]	7	Unknown	844	Against initial evaluation	Gleason groups[6]	85.3%	—
Berg et al, 2011 [4]	Unknown	Unknown	350	Against initial evaluation	Gleason groups[7]	76.9%	Substantial ($\kappa = 0.67$)
Kishimoto et al, 2012 [32]	1	Unknown	247	Against initial evaluation	Gleason groups[8]	61.9%	—
Goodman et al, 2012 [33]	3	Yes, to initial evaluation	268	Consensus score against initial evaluation	Gleason score Gleason groups[9]	—	*Gleason score* Moderate ($\kappa = 0.60$) *Gleason groups* Moderate ($\kappa = 0.58$)

(*continued*)

Table 4. Continued.

	Pathologists		Cases	Methods		Results	
Study	No.	Blinded	No.	Comparison standard	Outcome	Exact agreement	Interobserver concordance
Camara-Lopes et al, 2014 [19]	1	Unknown	182	Against initial evaluation	Gleason score	40.1%	—
Thomsen et al, 2015 [5]	1	Yes, to initial evaluation	109	Against initial evaluation	Gleason score	68.8%	Substantial ($\kappa_w = 0.67$)
Soga et al, 2015 [34]	Unknown	Unknown	388	Against initial evaluation	Gleason groups[6]	59.5%	—
Ozkan et al, 2016 [6]	2	Yes, to all other evaluations	407	Against each other	Gleason score ISUP grade	*Gleason score* 58% *ISUP grade* 51.7%	*Gleason score* Moderate ($\kappa = 0.43$) *ISUP grade* Fair ($\kappa = 0.39$)
Present study	Un-known	No	148	Against initial evaluation	ISUP grade	75%	Substantial ($\kappa_w = 0.76$)

[a]Gleason groups: 1 2–4, 5–6, 7, 8–10. 2 2–5, 6–7, 8–10. $3 < 7, \geq 7$. 4 2–4, 5–7, 8–10. 5 2–4, 5–6, 7–10. $6 \leq 6, 7, \geq 8$. $7 \leq 5, 6, 7, 8$–10. $8 < 5, 6, 3+4, 4+3, > 8$. $9 \leq 7, \geq 8$.
Abbreviation: No., number.

our study. A study of Nguyen et al. [35] included all patients who were referred to a genitourinary oncology specialist after being diagnosed with PCa and reported a change in risk group in 14.2% ($n = 92$) of the patients, which follows well with our finding.

To date, there are no recommendations in place in the Dutch and EAU guidelines on the use of pathologic re-evaluation of prostate biopsies in patients diagnosed with prostate cancer in the case of referral. In total, prostate biopsy re-evaluation was performed in 3% ($n = 172$) of the patients in our study. We showed that 93% ($n = 129$) of the prostate biopsy re-evaluations in patients treated with curative intent were performed in those who were referred to another hospital. For these patients the re-evaluation might have impacted treatment. However, the other 93% ($n = 1727$) of the patients treated with curative intent who were referred to another hospital did not have a pathology review. The conclusion is that, although re-evaluations might have important clinical implications, only in a small subset of patients is this review performed. Based on the current research, no firm recommendations on the necessity and cost-effectiveness of pathological review of prostate biopsies can be made yet.

A major strength of the current study is that we were able to provide insight into daily clinical practice by performing a nationwide population-based study. Most studies which assessed the interobserver concordance of the ISUP grade or EAU risk classification were single center studies performed in trial setting and are therefore not representative of daily clinical practice.

Conclusions

This population-based study shows that a histopathological re-evaluation of the prostate biopsies is infrequently requested by clinicians in the Netherlands. Despite the high interobserver concordance, risk classification changed as a result of the re-evaluation in one out of eight patients with localised PCa. This might have affected subsequent treatment and may thereby result in major therapeutic changes in a non-negligible minority of the patients. Future research should investigate the long-term clinical consequences and cost-effectiveness of re-evaluations in daily clinical practice before firm recommendations can be made.

Acknowledgements

The authors thank the registration team of the Netherlands Comprehensive Cancer Organisation for the collection of data for the Netherlands Cancer Registry.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Dutch Cancer Society under Grant [number IKNL 2013-5942].

References

[1] NKR cijfers [Internet]. Utrecht: IKNL; [cited 2019 Dec 15]. Available from: https://www.iknl.nl/nkr-cijfers/

[2] Mottet N, Bellmunt J, Bolla M, et al. EAU-ESTRO-SIOG guidelines on prostate cancer. Part 1: screening, diagnosis, and local treatment with curative intent. Eur Urol. 2017;71(4):618–629.

[3] Epstein JI, Allsbrook WC, Jr., Amin MB, et al. The 2005 International Society of Urological Pathology (ISUP) consensus conference on Gleason grading of prostatic carcinoma. Am J Surg Pathol. 2005;29(9):1228–1242.

[4] Berg KD, Toft BG, Roder MA, et al. Prostate needle biopsies: interobserver variation and clinical consequences of histopathological re-evaluation. APMIS. 2011;119(4–5):239–246.

[5] Thomsen FB, Marcussen N, Berg KD, et al. Repeated biopsies in patients with prostate cancer on active surveillance: clinical implications of interobserver variation in histopathological assessment. BJU Int. 2015;115(4):599–605.

[6] Ozkan TA, Eruyar AT, Cebeci OO, et al. Interobserver variability in Gleason histological grading of prostate cancer. Scand J Urol. 2016;50(6):420–424.

[7] Lancee M, Tikkinen KAO, de Reijke TM, et al. Guideline of guidelines: primary monotherapies for localised or locally advanced prostate cancer. BJU Int. 2018;122(4):535–548.

[8] Thomas CW, Bainbridge TC, Thomson TA, et al. Clinical impact of second pathology opinion: a longitudinal study of central genitourinary pathology review before prostate brachytherapy. Brachytherapy. 2007;6(2):135–141.

[9] Brimo F, Schultz L, Epstein JI. The value of mandatory second opinion pathology review of prostate needle biopsy interpretation before radical prostatectomy. J Urol. 2010;184(1):126–130.

[10] Sobin L, Gospodarowicz M, Wittekind C, et al. International Union Against Cancer (UICC) TNM classification of malignant tumors. 7th ed. Oxford (UK): Wiley-Blackwell; 2009.

[11] Epstein JI, Zelefsky MJ, Sjoberg DD, et al. A contemporary prostate cancer grading system: a validated alternative to the Gleason score. Eur Urol. 2016;69(3):428–435.

[12] Brennan PF, Hays BJ. The kappa statistic for establishing interrater reliability in the secondary analysis of qualitative clinical data. Res Nurs Health. 1992;15(2):153–158.

[13] Kweldam CF, Nieboer D, Algaba F, et al. Gleason grade 4 prostate adenocarcinoma patterns: an interobserver agreement study among genitourinary pathologists. Histopathology. 2016;69(3):441–449.

[14] Egevad L, Ahmad AS, Algaba F, et al. Standardization of Gleason grading among 337 European pathologists. Histopathology. 2013;62(2):247–256.

[15] Verhoef EI, van Cappellen WA, Slotman JA, et al. Three-dimensional analysis reveals two major architectural subgroups of prostate cancer growth patterns. Mod Pathol. 2019;32(7):1032–1041.

[16] Melia J, Moseley R, Ball RY, et al. A UK-based investigation of inter- and intra-observer reproducibility of Gleason grading of prostatic biopsies. Histopathology. 2006;48(6):644–654.

[17] Egevad L, Algaba F, Berney DM, et al. Interactive digital slides with heat maps: a novel method to improve the reproducibility of Gleason grading. Virchows Arch. 2011;459(2):175–182.

[18] Rider JR, Sandin F, Andren O, et al. Long-term outcomes among noncuratively treated men according to prostate cancer risk category in a nationwide, population-based study. Eur Urol. 2013;63(1):88–96.

[19] Camara-Lopes G, Marta GN, Leite ET, et al. Change in the risk stratification of prostate cancer after Slide Review by a uropathologist: the experience of a reference center for the treatment of prostate cancer. Int Braz J Urol. 2014;40(4):454–459. discussion 460–2.

[20] Steinberg DM, Sauvageot J, Piantadosi S, et al. Correlation of prostate needle biopsy and radical prostatectomy Gleason grade in academic and community settings. Am J Surg Pathol. 1997;21(5):566–576.

[21] Lessells AM, Burnett RA, Howatson SR, et al. Observer variability in the histopathological reporting of needle biopsy specimens of the prostate. Hum Pathol. 1997;28(6):646–649.

[22] Wurzer JC, Al-Saleem TI, Hanlon AL, et al. Histopathologic review of prostate biopsies from patients referred to a comprehensive cancer center: correlation of pathologic findings, analysis of cost, and impact on treatment. Cancer. 1998;83(4):753–759.

[23] Allsbrook WC, Jr., Mangold KA, Johnson MH, et al. Interobserver reproducibility of Gleason grading of prostatic carcinoma: urologic pathologists. Hum Pathol. 2001;32(1):74–80.

[24] Renshaw AA, Schultz D, Cote K, et al. Accurate Gleason grading of prostatic adenocarcinoma in prostate needle biopsies by general pathologists. Arch Pathol Lab Med. 2003;127(8):1007–1008.

[25] Glaessgen A, Hamberg H, Pihl CG, et al. Interobserver reproducibility of percent Gleason grade 4/5 in prostate biopsies. J Urol. 2004;171(2 Pt 1):664–667.

[26] Coard KC, Freeman VL. Gleason grading of prostate cancer: level of concordance between pathologists at the University Hospital of the West Indies. Am J Clin Pathol. 2004;122(3):373–376.

[27] Oyama T, Allsbrook WC, Jr., Kurokawa K, et al. A comparison of interobserver reproducibility of Gleason grading of prostatic carcinoma in Japan and the United States. Arch Pathol Lab Med. 2005;129(8):1004–1010.

[28] Sooriakumaran P, Lovell DP, Henderson A, et al. Gleason scoring varies among pathologists and this affects clinical risk in patients with prostate cancer. Clin Oncol (R Coll Radiol)). 2005;17(8):655–658.

[29] Griffiths DF, Melia J, McWilliam LJ, et al. A study of Gleason score interpretation in different groups of UK pathologists; techniques for improving reproducibility. Histopathology. 2006;48(6):655–662.

[30] Fine SW, Epstein JI. A contemporary study correlating prostate needle biopsy and radical prostatectomy Gleason score. J Urol. 2008;179(4):1335–1338.

[31] Veloso SG, Lima MF, Salles PG, et al. Interobserver agreement of Gleason score and modified Gleason score in needle biopsy and in surgical specimen of prostate cancer. Int Braz J Urol. 2007;33(5):639–646.

[32] Kishimoto R, Saika T, Bekku K, et al. The clinical impact of pathological review on selection the treatment modality for localized prostate cancer in candidates for brachytherapy monotherapy. World J Urol. 2012;30(3):375–378.

[33] Goodman M, Ward KC, Osunkoya AO, et al. Frequency and determinants of disagreement and error in gleason scores: a population-based study of prostate cancer. Prostate. 2012;72(13):1389–1398.

[34] Soga N, Yatabe Y, Kageyama T, et al. Review of bioptic gleason scores by central pathologist modifies the risk classification in prostate cancer. Urol Int. 2015;95(4):452–456.

[35] Nguyen PL, Schultz D, Renshaw AA, et al. The impact of pathology review on treatment recommendations for patients with adenocarcinoma of the prostate. Urol Oncol. 2004;22(4):295–299.