Interobserver reproducibility in defining morphological parameters of patients with non-muscle-invasive bladder cancer with poor prognosis

T.I. Nabebina, A.I. Rolevich, A.Ch. Dubrovsky, S.A. Krasny

N.N. Aleksandrov Republican Research and Practical Center for Oncology and Medical Radiology; Lesnoy settlement, Minsk region, 223 040, Republic of Belarus

Contacts: *Tatyana Ivanovna Nabebina nabebina.t@yandex.by*

Purpose – to evaluate the interobserver variability in grading and staging of non-muscle-invasive bladder urothelial carcinoma (UC) and to determine the prognostic value of data obtained from the second pathology review. Light microscopy of archived samples of 158 patients was redone during the study. We used statistical parameters in comparing the initial pathology conclusions with the data obtained after the second pathology review. We found high variability of histological conclusions as well as low interobserver agreement in defining morphological parameters. We revealed a tendency toward the higher UC grade after the second pathology review, pT1 category was confirmed in 40% patients and it was down staged to pTa in 60% cases.

Key words: *bladder cancer, grade, pT category, interobserver variability, prognostic value*

DOI: 10.17650/1726-9776-2016-12-1-21-28

Introduction

Urothelial carcinoma of the bladder (UCB), also known as transitional cell carcinoma (TCC) of the bladder, comprised of 2,7 % of all malignancies in the Republic of Belarus in 2013, accounting for 4,4 % of this cancer in males and 1,1 % in females. The incidence rate for TCC per 100.000 people was approximately 15,2 cases for males and 2,1 for females in 2009–2013 years [1]. Bladder cancer (BC) is the fifth most common malignancy in Western countries with the highest lifetime treatment costs per patient of all cancers [2]. Approximately 75–85 % of all patients with BC are initially diagnosed with «non-muscle-invasive disease» that is confined to the mucosa (pTa and carcinoma in situ – 70 % and 10 % of all cases, respectively) and submucosa (pT1–20 %) [3]. These tumors represent a heterogenous group of malignancies with highly variable recurrence rates and progression up to muscle invasion. According to statistics, the risk of recurrence for TCC is from 15 % to 70 % within the first year after treatment, with the risk of further progression up to muscle invasion from 7 to 40 % within the first 5 years [3, 4].

Among the main factors that determine the aggressive behavior of non-muscle-invasive UCB is tumor morphology, which includes the grade (the 1973 WHO classification (G1-G3)), pT category (pTa, pT1), concomitant carcinoma in situ (CIS) [4]. There are significant differences in the definition of the tumor stage and grade when it is done initially and after the second pathology review by different specialists (interobserver variability) as well as by the same doctor (intraobserver variability). As tumor morphology is considered to be the main factor that determine the treatment strategy (repeat transurethral resection, intravesical immunotherapy with BCG, cystectomy), the reproducibility of morphological results, as well as the fact of the second pathology review can significantly affect the therapeutic strategy and patient outcomes.

Therefore, the purpose of our study was to evaluate the reproducibility in grading and staging of non-muscle-invasive UCB, which was done by several specialists and to determine the prognostic value of data obtained from the second pathology review.

Materials and methods

We studied a group of patients with non-muscle-invasive BC. All patients were after radical transurethral resection which was performed in the urology department of the RRPC OMR named after N.N. Alexandrov in 2004–2007. Some of them were treated with adjuvant intravesical BCG immunotherapy. We analyzed a range of characteristics which were suggested by EORTC (European Organization for Research and Treatment of Cancer) as the main criteria for predicting recurrence and progression up to muscle invasion. Among them were tumor grade (the 1973 WHO classification (G1 – G3)), pT category (pTa, pT1), concomitant CIS, multifocal disease (solitary, 2–7 tumors), tumor size (< 3 cm, ≥ 3 cm), the first recurrence time (primary tumor, <1 recurrence per year, >1 recurrence per year). Tumor grade, pT category and concomitant CIS are the most important prognostic factors for tumor progression with lamina propria invasion [4, 5]. We evaluated the progression risk index for every patient according to the presence and combination of parameters. We selected 158 patients with an estimated risk of progression ≥7 (unfavorable prognosis).

We redid the light microscopy of hematoxylin and eosin stained archived samples of 158 patients with TCC. The samples were reviewed by two experienced pathologists who participated in the study. In case of interobserver disagreement, the sample was viewed on one more microscope and the final conclusion was done only after an interobserver agreement was reached. Samples of 4 patients were excluded from further analysis due to the insufficient quantity and poor quality of the material.

Grade assessment of TCC was done according to the 1973 [6] and the 2004 [7] WHO classification systems as well as with the use of diagnostic algorithm for papillary urothelial tumors [8]. According to this algorithm each parameter is given individual scores (from 0 to 3 for mitosis and cellular thickness, from 0 to 2 for cellular atypia, and an additional score for papillary fusion). These scores were combined to form a summed score allowing the tumors to be ranked as follows: 0–1 – urothelial papilloma, 2–4– low malignant potential carcinoma, 5–7 – low-grade TCC and 8–9 – high-grade TCC (table 1, 2) [8].

Table 1. *Diagnostic scoring scheme for papillary transitional cell carcinoma grading*

Parameter	Definition	Score
Mitosis	No < 5 per 10 high power field, ×40 5–10 > 10	0 1 2 3
Thickness of layers	< 7 layers with intact umbrella cells < 7 layers with the loss of umbrella cells > 7 layers with intact umbrella cells >7 layers with the loss of umbrella cells	0 1 2 3
Cellular atypia	No Diffuse in a random fashion, but with a mild degree Diffuse in a random fashion, but with a moderate to severe degree	0 1 2
Papillary fusion	No Yes	0 1

Submucosal invasion was defined as the prevalence of isolated tumor cells and variably sized nests with small clusters of tumor cells irregularly invading submucosa with fibrosis reaction of the stroma [9]. For the T1 substage, we used a system that discerns T1-microinvasive (pT1m) and T1-extensive-invasive (T1e) tumors [10]. If we found a solitary submucosal lesion of ≤0,5 mm (per 1 high power field, x 40), the tumor was defined as pT1m. If we found a solitary submucosal lesion of >0,5 mm or multiple microinvasive submucosal lesions, the tumor was defined as pT1e (fig. 1, 2).

We compared the initial pathology conclusions with the data obtained after the second pathology review, estimated the percent agreement figure, Cohen»s kappa coefficient and its 95 % confidence intervals (CI). We assessed progression free survival according to the grade and pT category which were known initially and after the second pathology review.

Results

Median age of patients with non-muscle-invasive BC was 64 (29-92) years. Incidence rates were higher in males compared to females with the male-female ratio of 3,2:1. Primary tumor was diagnosed in 100 (64,5 %) patients and 55 (35,5 %) patients had recurrent disease. Multifocal disease was found by cystoscopy in a majority of patients (n = 126; 81 %). The median number of lesions was 3 (2-21). Papillary tumor with a broad base was the most common pattern of tumor growth (n = 144; 93 %). Tumor size varied from 0,3 to 7,0 cm (median 2,5 cm). The main baseline characteristics of patients as well as cystoscopy findings are depicted in table 3.

Grade assessment was initially done in 128 (83,1 %) patients according to the 1973 WHO classification and in 26 (16,8 %) cases it was not specified. Tumor grading on well differentiated (G1), moderately differentiated (G2) and poorly differentiated (G3) tumors was done in 75 (58,6 %), 41 (32,0 %) and 12 (9,4 %) patients, respectively. Only 62 (48,4 %) patients among them had the same tumor grade after the second pathology review; in 52 (40,6 %) cases it was upstaged and in 14 (10,9 %) cases it was down staged. Furthermore, in 6 (4,7 %) cases the defined tumor grade was

Table 2. *Histologic grading according to the 1998 WHO/ISUP classification*

Neoplasms	Mitosis	Thickness of layers	Atypia	Papillary fusion	Score
UP – urothelial papilloma	0–1	0	0	0	0–1
IP – inverted papilloma	0–1	0–1	0–1	1	1–3
PUNLMP – papillary urothelial neoplasm of low malignant potential	0–1	1–2	0	1	2–4
LGPUC– low-grade papillary urothelial carcinoma	2	1–3	1	1	5–7
HGPUC – high-grade papillary urothelial carcinoma	3	2–3	2	1	8–9

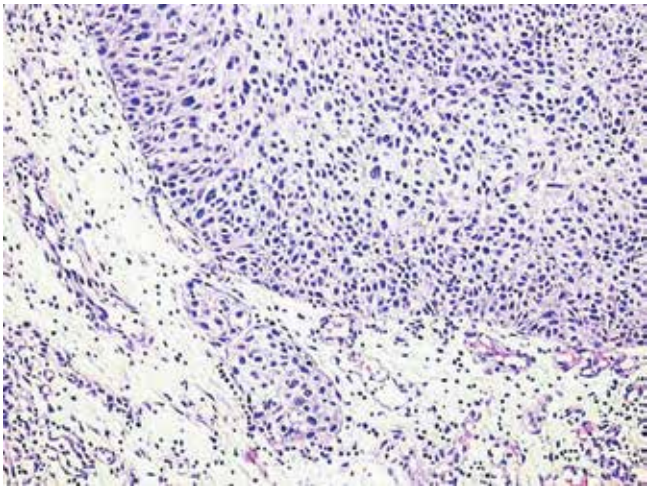

Fig. 1. *Submucosal invasion, pT1m. Hematoxylin and eosin staining, ×10*

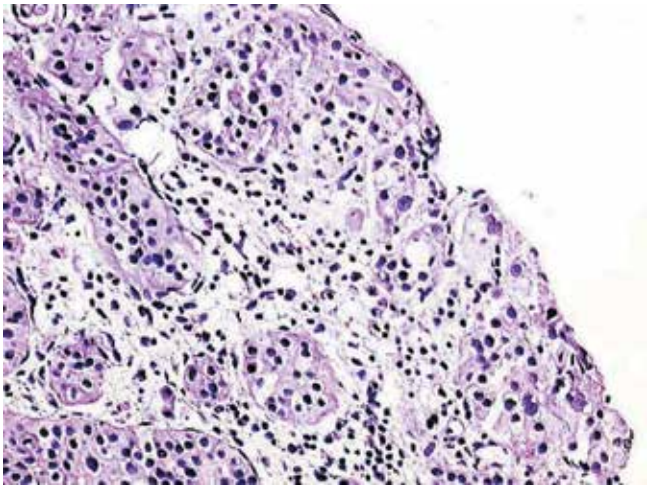

Fig. 2. *Submucosal invasion, pT1e. Hematoxylin and eosin staining, ×20*

changed by more than 1 grade. Grade assessment was initially done in 124 (80,5 %) patients according to the 2004 WHO classification and in 30 (19,4 %) cases it was not specified. Low-grade tumors as well as high-grade tumors were defined in 103 (83,1 %) and 21 (16,9 %) patients, respectively. Only 66 (53,2 %) patients among them had the same tumor grade after the second pathology review; in 32 (25,8 %) cases it was upstaged and in 26 (20,9 %) cases it was down staged. The degree of agreement between the 1973 and the 2004 WHO classifications is depicted in table 4.

pTa stage was defined initially in 1 (0,6 %) case and pT1 stage – in 153 (99,4 %) patients. After the second pathology review these stages were defined in 93 (60,0 %) and 62 (40,0 %) patients, respectively. While in 1 patient with the initial pTa stage the diagnosis was confirmed after the pathology review, only 62/153 (40,0 %) patients with pT1 stage had the same diagnosis and in 92/153 (60 %) cases it was down staged after the review. Distribution of patients as per the initial pT stage and results after the pathology revision is depicted in table 5. In comparison to initial conclusion, histological review led to the increase of the grade and decrease of the pT category (fig. 3).

According to our results, there was a low interobserver agreement in UC grade, which was independent of the grading system (the 1973 and the 2004 WHO classification). The accepted standard enabled to divide this agreement into slight agreement, slight (ns) agreement and fair agreement. The major agreement (moderate agreement) was reached in defining poorly differentiated TCC, the type which is primarily characterized by the loss of structural organization and anaplasia. We found no interobserver agreement (no agreement NS)) in defining pT category. Interobserver agreement, done by different pathologists with the use of the 1973 and the 2004 WHO classification systems, as well as separate subgroups within these classifications and pT category are depicted in table 5.

Table 3. *Baseline characteristics of patients (median age 64 (29–92) years) and cystoscopy findings*

Characteristic	*n*	%
Gender		
females	37	24
males	118	76
Tumor		
primary	100	65
recurrent	55	35
Multifocal disease		
solitary	29	19
2–7 lesions	96	62
≥ 8 lesions	30	19
Macroscopic view of the tumor		
papillary	144	93
solid/ulcerating tumor	11	7
Maximum tumor size		
< 3 cm	81	52
> 3 cm	69	45
No data	5	3

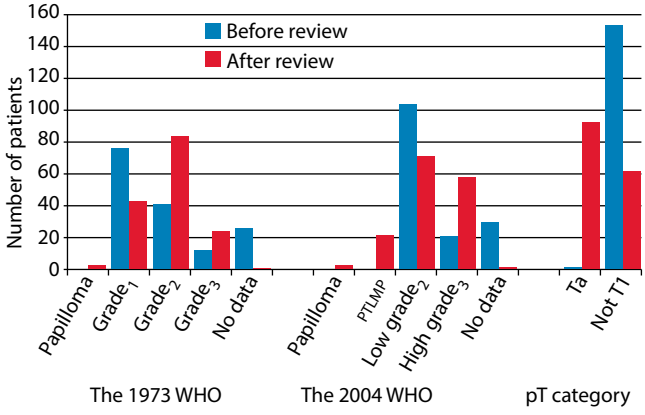

Fig. 3. *Comparison of initial conclusion with the results of pathology review*

Table 4. *The degree of agreement in grade (the 1973 and the 2004 WHO classification systems) and pT category as per the time of initial diagnosis and after the second pathology review*

Grade assessment according to the 1973 WHO classification, *n* (%)

Initial findings	Findings after the pathology review				Total
	Papilloma	G_1	G_2	G_3	
G_1	3 (4)	25 (33)	42 (56)	5 (7)	75 (100)
G_2	0	8 (20)	28 (68)	5 (12)	41 (100)
G_3	0	1 (8)	2 (17)	9 (75)	12 (100)

Grade assessment according to the 2004 WHO classification, *n (%)*

Initial findingse	Findings after the pathology review				Total
	Papilloma	PTLMP	Low grade	High grade	
Low grade	3 (3)	18 (17)	50 (49)	32 (31)	103 (100)
High grade	0	1 (5)	4 (19)	16 (76)	21 (100)

pT category

Initial findings	Findings after the pathology review		Total
	Ta	T1	
Ta	1 (100)	0	1 (100)
T1	92 (60)	62 (40)	154 (100)

Note: *PTLMP – papillary tumor of low malignant potential*

Table 5. *The degree of agreement between the initial findings and findings after the pathology review*

Characteristic	Percent agreement (95 % CI)	Cohen's kappa coefficient (95 % CI)	Notes
The 1973 WHO	48 (40–57)	0,21 (0,08–0,33)	Fair agreement
G_1	54 (45–63)	0,15 (0,01–0,28)	Slight agreement
G_2	55 (47–64)	0,15 (–0,002–0,30)	Slight (NS) agreement
G_3	90 (85–95)	0,53 (0,30–0,75)	Moderate agreement
The 2004 WHO	53 (44–62)	0,18 (0,07–0,29)	Slight agreement
Low-grade	54 (46–64)	0,15 (0,04–0,27)	Slight agreemen
High-grade	70 (62–78)	0,30 (0,14–0,46)	Fair agreement
pT category	41 (33–48)	0,01 (–0,01–0,03)	No agreement (NS)

Median follow-up of all patients included in our study was 76 (4–102) months. Progression free survival according to the grade (G) and pT category defined primarily and after the second pathology review is depicted in figure 4. All patients were divided into the groups by the tumor grade with statistically significant difference in progression free survival rates at the time of initial diagnosis and after the review. Histological revision enabled to determine subgroups of patients with favorable prognosis (urothelial papilloma and urothelial neoplasm of low malignant potential), as well as to divide well-differentiated and moderately differentiated TCC by the risk of progression (according to the 1973 WHO). It was impossible to say whether initially established pT category refers to the prognosis of the disease or not, as the majority of cases were defined as pT1. Further revision of pT category enabled to improve the distribution of patients according to prognostic groups. However, we did not reach statistically signifi-

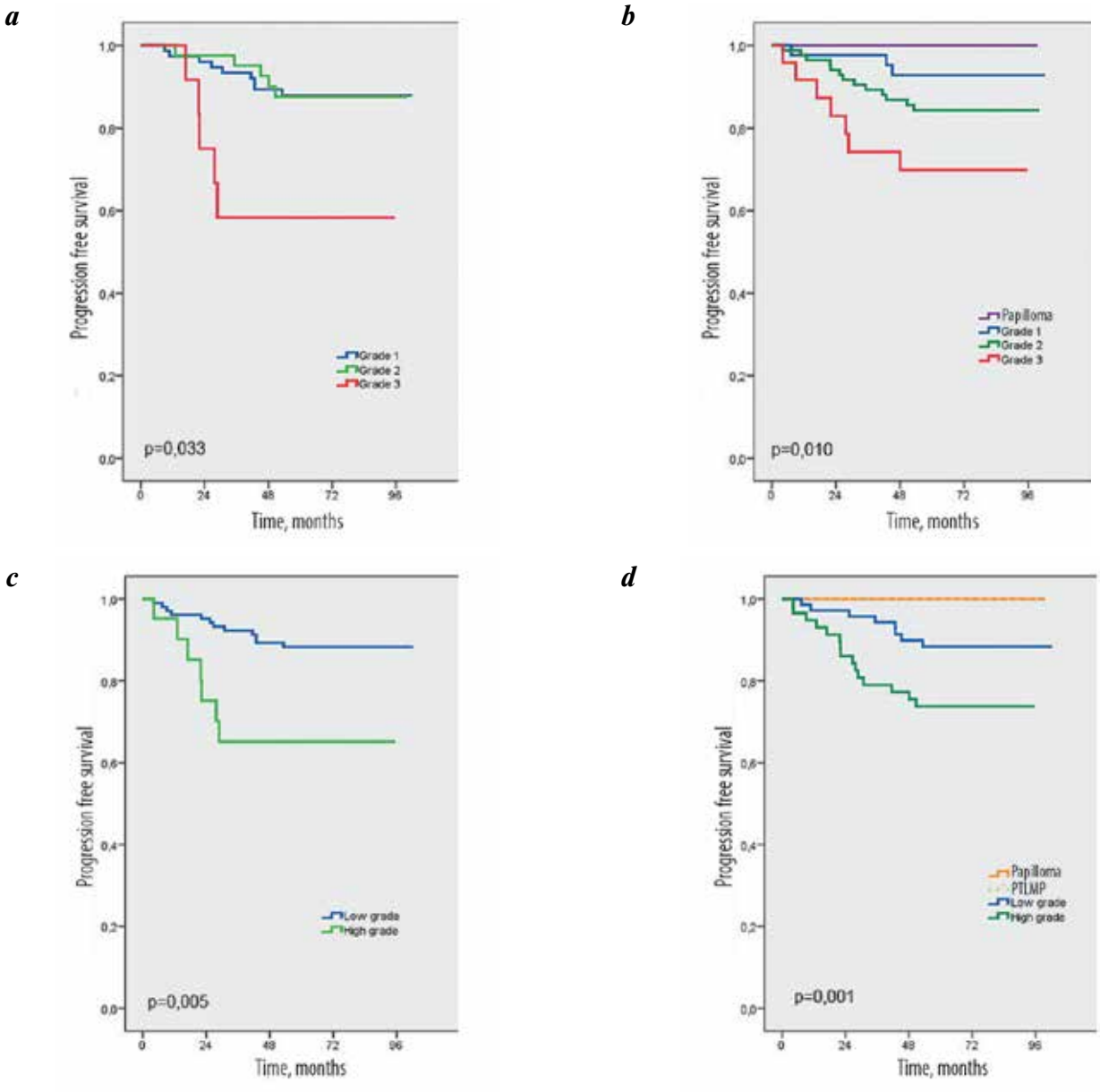

Fig. 4. *Progression free survival according to the grade defined initially and after the second pathology review (the 1973 and the 2004 WHO classification systems): a - the 1973 WHO, initial diagnosis; b – the 1973 WHO, review; c - the 2004 WHO, initial diagnosis; d – the 2004 WHO, review*

cant meanings in the difference of progression free survival among pTa and pT1 groups (fig. 5).

Discussion

UCB is a complex and heterogenous disease with a potentially lethal behavior. One of the main factors that determine the treatment strategy as well as the risk of recurrence up to muscle invasion is tumor morphology. There are significant differences in the definition of the tumor stage and grade (level of invasion) when it is done initially and after the second pathology review by different specialists (interobserver variability) as well as by the same doctor (intraobserver variability). The existence of variability in histological evaluation is proved in a variety of studies. According to the results of one randomized multicentre trial, which was conducted by Witjes JA and colleagues [11], the conformity between local and review pathology of the pT category was 79,3 %, of the grade 70,2 %, and the combination of both 59,7 %. In local pathology, undergrading was more frequent than overgrading and overstaging more frequent than understaging. However, the authors noticed that the prognostic relevance of tumor grade remained the same after correction for review pathology. Only the prognostic relevance of tumor stage increased after pathology correction.

In the trial conducted by Van Der Meijden A and colleagues [12] local and review pathology results were

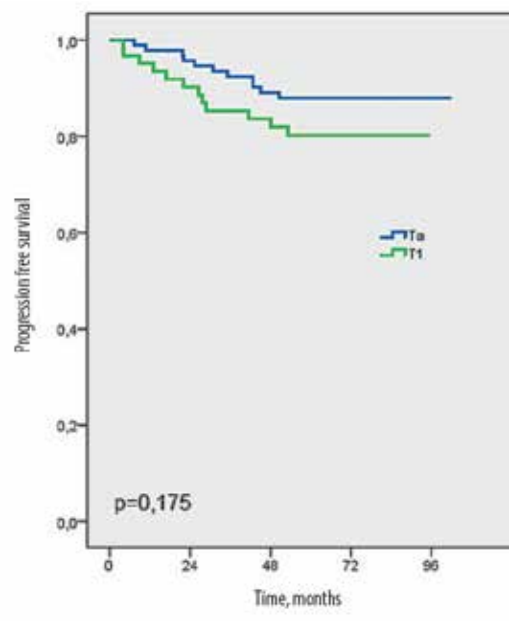

Fig. 5. *Progression free survival according to the pT category defined after the second pathology review*

assessed in 1400 patients with primary or recurrent urothelial non-muscle-invasive carcinoma of the bladder. The authors found large interobserver variability in T category and grade. Pathology review down staged T category to stage Ta in 53 % of cases originally classified as stage T1. There was agreement in only 57 % and 50 % of stage Ta grade 1 and stage T1 grade 3 cases. While T category and grade have prognostic importance, differences in the prognosis based on local and review pathological studies were slight. However, the authors highly recommend to do pathology review in high risk cases of stage T1 grade 3 disease as 10 % of them were reclassified as muscle invasive disease (greater than stage T1). The last one led to certain changes in the treatment decision process.

In the trial conducted by Bol MG and colleagues the consensus and original diagnoses agreed on stage and grade initially and after the second review in 68,5 % and 62,3 % of cases, respectively [13]. Of the original diagnoses of stage T1 tumors the consensus diagnosis down staged 55,6 % to Ta and up staged 12,7 % to T2–3. Staging and pT grading defined after the second review had the highest prognostic significance. In the trial conducted by Tosoni I and colleagues [14] there were significant interobserver differences in both the grading and staging of tumours. From a total of 235 tumours that were initially considered pT1, the reviewer classified 35 % as pTa, 56 % as pT1, 6 % as pT1- (at least pT1), and 3 % as pT2–4. In 39 % of all biopsies there were interobserver differences in tumour grade. The reviewer»s staging allowed a better estimate of the risk of subsequent tumour progression than the initial staging. From the trial conducted by van Rhijn BW and colleagues [15] it was concluded that a stage review is indicated in pT1 BC, as almost 20 % of pT1 tumours were up- or down-staged, and the reviewed stage predicted the patient»s prognosis.

Despite provision of detailed histologic criteria for the diagnostic categories in the 2004 WHO system, improvement in intraobserver and interobserver variability as compared to the 1973 WHO system has not been documented [16–19]. Mikuz G et al [20] demonstrated that interobserver agreement was higher using the 1973 WHO classification than when using either the 2004 WHO or 1999 WHO/ISUP systems. In a study by Yorukoglu K and colleagues [21], no statistical difference between the intraobserver and interobserver reproducibility of both the 2004 WHO and the 1973 WHO systems was achieved; the new system failed to improve reproducibility. Therefore some scientists think that the WHO classification of 1973 is more suitable than the WHO classification of 2004 for predicting survival in patients with UCB [22].

Conclusion

Our study confirmed the existence of high variability of histological conclusions as well as low inter-observer agreement in UC grade. Revision of histological samples confirmed the same grade in 48,4 % and 53,2 % patients (according to the 1973 and the 2004 WHO classifications, respectively). We revealed a tendency toward the higher tumor grade in 40,6 % and 25,8 % of cases (according to the 1973 and the 2004 WHO classifications, respectively). In a majority of cases we found low interobserver agreement with moderate agreement defined in poorly differentiated UCB. Both classifications (the 1973 and the 2004 WHO) equally stratified patients into risk groups for disease progression at the time of initial diagnosis and after the second pathology review. Histological revision enabled to determine subgroups of patients with favorable prognosis (urothelial papilloma and urothelial neoplasm of low malignant potential) as well as to divide groups of well-differentiated (G1) and moderately differentiated (G2) tumors according to the 1973 WHO classification.

We did not find any interobserver agreement in UC pT staging. Initially established pT1 category was confirmed in 62 (40 %) patients and it was down staged to pTa in 92 (60 %) cases. According to the data obtained after the second pathology review, groups of patients with pT1 stage had worse survival. However, this result was not statistically significant that is probably could be explained by the low statistical power of our trial.

According to our results, evaluation of UC grade as well as pT category by different pathologists have unequal prognostic potential (it is mainly relevant to pT category). Therefore, if only it is possible you»ll have to receive histological conclusions done by two independent pathologists when you think about radical treatment of patients with BC.

REFERENCES

1. Okeanov A.E., Moiseyev P.I., Levin L.F. Statistics of cancer diseases in the Republic of Belarus (2004–2013). Edited by Sukonko O.G. Minsk: N.N. Alexandrov RRPC OMR, 2014. P. 171–177 (In Russ.).
2. Sievert KD, Amend B., Nagele U. et al. Economic aspects of bladder cancer: what are the benefits and costs? World J Urol 2009;27:295–300.
3. van Rhijn BWG, Burger M., Lotan Y. et al. Recurrence and progression of disease in nonmuscleinvasive bladder cancer: from epidemiology to treatment strategy. EurUrol 2009;56:430–42.
4. Sylvester RJ., vander Meijden APM., Oosterlinck W. et al. Predicting recurrence and progression in individual patients with stage Ta T1 bladder cancer using EORTC risk tables: a combined analysis of 2596 patients from seven EORTC trials. EurUrol 2006;49:466–77.
5. FernandezGomez J., Solsona E., Unda M. et al. Prognostic factors in patients with non–muscleinvasive bladder cancer treated with bacillus CalmetteGue′rin: multivariate analysis of data from four randomized CUETO trials. EurUrol 2008;53:992–1002.
6. Mostofi FK., Sobin .H, Torloni H. Histological typing of urinary bladder tumours. Geneva, Switzerland: World Health Organization, 1973.
7. Eble J., Sauter G., Epstein J. et al. Pathology and genetics of tumours of the urinary system and male genital organs. Lyon, France: IARC Press, 2004.
8. Shim J.W., Cho KS., Choi YD. et al. Diagnostic algorithm for papillary urothelial tumors in the urinary bladder. VirchowsArch 2008;452:353–62.
9. Bostwick DG., Mikuz G. Urothelial papillary (exophytic) neoplasms. VirchowsArch 2002;44:109–16.
10. Van Rhijn BWG., vander Kwast TH., Alkhateeb SS. et al. A new and highly prognostic system to discern T1 bladder cancer substage. EurUrol 2012;61:378–84.
11. Witjes JA., Kiemeney LALM., Schaafsmas HE. et al. The influence of review pathology on study outcome of a randomized multicentre superficial bladder cancer trial. BJU 1994;73:172–6.
12. VanderMeijden A., Sylvester R., Collette L. et al. The role and impact of pathology review on stage and grade assessment of stagesTa and T1 bladder tumors: a combined analysis of 5 european organization for research and treatment of cancer trials. J Urol 2000;164: 1533–7.
13. Bol MGW., Baak JPA., BuhrWildhagen S. et al. Reproducibility and prognostic variability of grade and lamina propria invasion in stages Ta, T1 urothelial carcinoma of the bladder. J Urology 2003;169: 1291–4.
14. Tosoni I., Wagner U., Sauter G. et al. Clinical significance of interobserver differences in the staging and grading of superficial bladder cancer. BJU International 2000;85:48–53.
15. Van Rhijn BWG., vander Kwast TH., Kakiashvili TH. et al. Pathological stage review is indicated in primary pT1 bladder cancer. BJU Int 2009;106:206–11.
16. Engers R. Reproducibility and reliability of tumor grading in urological neoplasms. World J Urol 2007;25:595–605.
17. May M., BrookmanAmissah S., Roigas J. et al. Prognostic accuracy of individual uropathologists in noninvasive urinary bladder carcinoma: a multicentre study comparing the 1973 and 2004 World Health Organisation classifications. EurUrol 2010;57:850–8.
18. Van Rhijn BWG., van Leenders GJ., Ooms BC. et al. The pathologist's mean grade is constant and individualizes the prognostic value of bladder cancer grading. EurUrol 2010;57:1052–7.
19. Gonul II, Poyraz A., Unsal C. et al. Comparison of 1998 WHO/ISUP and 1973 WHO classifications for interobserver variability in grading of papillary urothelial neoplasms of the bladder. Pathological evaluation of 258 cases. UrolInt 2007;78:338–44.
20. Mikuz G. The reliability and reproducibility of the different classifications of bladder cancer. In: Haupmann S., Dietel M., SorbrinhoSimoes M., edit. Surgical pathology update 2001. Berlin: ABW-Wissenschaftsverlag, 2001. P. 114–115.
21. Yorukoglu K., Tuna B., Dikicioglu E. et al. Reproducibility of the 1998 World Health Organization/International Society of Urologic Pathology classification of papillary urothelial neoplasms of the urinary Bladder. VirchowsArch 2003;443:734–40.
22. Otto W., Denzinger S., Fritsche HM. et al. The WHO classification of 1973 is more suitable than the WHO classification of 2004 for predicting survival in pT1 urothelial bladder cancer. BJU Int 2010;107:404–8.