



# Comparison of Interobserver Variability Between Two Grading Systems Used for Urothelial Carcinoma

Shirin Taraz Jamshidi <sup>1</sup>, Mohaddeseh Shahraki <sup>1,\*</sup>, Mahmoud Reza Kalantari <sup>1</sup>, Mehdi Farzadnia <sup>2</sup>, Amirhossein Jafarian <sup>2</sup>, Negar Morovatdar <sup>3</sup>

<sup>1</sup> Department of Pathology, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran

<sup>2</sup> Cancer Molecular Pathology Research Center, Mashhad University of Medical Sciences, Mashhad, Iran

<sup>3</sup> Clinical Research Developmental Unit, Imam Reza Hospital, Mashhad University of Medical Sciences, Mashhad, Iran

\*Corresponding author: Department of Pathology, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran. Email: m.shahraki1991@gmail.com

Received 2023 December 9; Revised 2024 May 11; Accepted 2024 June 7.

## Abstract

**Background:** Inter-observer variability among pathologists' reports has been a significant challenge in the diagnosis of diseases, especially cancers. Designing a grading system with minimal inter-observer variability can help achieve more accurate diagnoses.

**Objectives:** We aimed to determine whether the grading system proposed by Cheng and colleagues (Cheng's grading system) can improve inter-observer variability among pathologists in grading urothelial cancer, compared with the World Health Organization (WHO) grading system (2004/2016).

**Methods:** Four pathologists examined all slides of bladder biopsy samples diagnosed as urothelial carcinoma, available in the archives of Imam Reza Hospital, Mashhad, Iran, from 2019 to 2022. Each pathologist reported the tumor grade based on both grading systems, independently of previous answers or the reports of their colleagues.

**Results:** Of 132 samples, the majority were from men (84.1%); the mean age of patients was 66.56±11.31 years. For low and high grades of the WHO system and grade II in Cheng's grading system,  $\kappa = 0.602$ ; agreement was lower for grade III ( $\kappa = 0.439$ ) and higher for grade IV ( $\kappa = 0.690$ ) compared with the WHO grading system. There was significant agreement among the pathologists in all tumor characteristics, with excellent agreement for stromal invasion ( $\kappa = 0.790$ ) and muscular invasion ( $\kappa = 0.884$ ), fair to good agreement for WHO 2004/2016 grading ( $\kappa = 0.602$ ) and Cheng's grading ( $\kappa = 0.574$ ) overall, and poor agreement for tumor heterogeneity ( $\kappa = 0.360$ ).

**Conclusions:** Both grading systems demonstrated favorable agreement among pathologists, without significant differences between the two systems. These results confirm the appropriateness of both grading systems.

**Keywords:** Classification, Carcinoma, Observer Variation, Urinary Bladder Neoplasms

## 1. Background

Urothelial cancer of the urinary bladder (UCB) is the most common type of bladder cancer (90%) and the fourth most common cancer globally, with a male-to-female ratio of 2: 1 (1, 2). The increasing trend in its incidence, high mortality rate, and poor prognosis (overall survival of 12 - 14 months) impose a substantial public health burden (3).

Early diagnosis of the histological type, muscular and non-muscular invasion, and histological characteristics of the tumor can help improve the prognosis of this

cancer. Accordingly, several grading systems have been suggested for the classification of UCB; one of the most widely accepted is the one proposed by the World Health Organization (WHO). After several modifications (1973, 1998, and 2004) (4), a clear classification was proposed in 2016 for each grade, which avoided ambiguous classifications and differentiated non-invasive and invasive carcinoma (5-7). Another classification is that proposed by Cheng et al. which classified UCB into four grades: I and II (low grades), and III and IV (high grades), and abandoned using other terms (8).

Therefore, it is not specified which is the best grading system for UCB; some pathologists prefer to report results based on the two most recent grading systems (9). However, it is unclear which of these two grading systems, namely WHO 2004/2016 and Cheng's grading system, is superior. Previous studies have compared different versions of WHO grading systems but have not compared them with Cheng's system.

Reproducible results, obtained by low inter- and intra-observer variability among pathologists, are important goals of classification systems. After the "2004 WHO system" was proposed, several studies compared its variable agreement with previous systems (10). Its comparison with the "1973 WHO system" reported marginally better reproducibility for the WHO 2004/2016 system (11-13). Meanwhile, others suggested better agreement with the 1973 version compared with the later versions, 2004 and 1999 (14), and some have reported poor reproducibility for both the 2004 and 1973 classifications (15). One limitation of these studies was disregarding the agreement between pairs of pathologists (13). Another limitation was the consideration of only specific tumors, such as TaT1 carcinoma (16) or non-invasive carcinoma (12); therefore, their results are not generalizable to all UCBs.

Tumor heterogeneity is an important source of variability, especially in tumor grade III, as it has features of both low and high grades, while grade IV is typically high-grade, and grade II is typically low-grade. This is why higher agreement has been reported in low grades (I and II; 89%) compared to grade III (66%) (16). Cheng and colleagues have suggested specifying grade III as a distinct grade (8). Therefore, we hypothesized that Cheng's system may be as accurate as the 2004/2016 WHO system or even superior to it.

## 2. Objectives

As previous studies have not compared the reproducibility of the results reported by these two systems, and considering the advantages counted for this new classification system (Cheng's system), this study aimed to demonstrate whether Cheng's grading system can improve the inter-observer variability among pathologists for staging urothelial cancer compared with the 2004/2016 WHO grading system.

## 3. Methods

All slides of bladder biopsy samples diagnosed as urothelial carcinoma, available in the archives of Imam Reza Hospital, Mashhad, Iran (related to patients who were referred from April 21, 2019, to March 22, 2022),

were collected using a convenience sampling method. As treatment can cause changes in urothelial cells, mimicking dysplasia and carcinoma (17), we excluded cases who had received treatment (N = 3). Four expert pathologists from Mashhad, including one uropathologist and three general pathologists, two from Imam Reza Hospital and two others from Qaem Hospital, examined all slides. All pathologists were board-certified, members of Mashhad University of Medical Sciences, with 10 - 21 years of experience, and working in university-affiliated hospitals. In cases where the slide was not available or was of low quality, two slides were prepared from the relevant paraffin block.

Each pathologist reported the tumor grading based on the two grading systems, 2004/2016 WHO and Cheng et al.'s grading system, as well as the tumor characteristics (stromal invasion, muscular invasion, and tumor heterogeneity). Heterogeneity was considered when the tumor's grade was mixed (also named as the mixed-grade tumor), and different histological grades (low grade and high grade) were observed in different areas of the specimen evaluated (18, 19). Grades in Cheng's proposal were defined as I, II, III, and IV; grades I and II were considered low-grade, and grades III and IV high-grade (8). Comparing Cheng's system to the WHO 2004/2016 system, grade I is equivalent to PUNLMP, grade II to low grade, and grades III/IV to high grade (8). The pathologists were not aware of the previous answers or the reports of other colleagues. The slides not diagnosed as urothelial carcinoma were excluded from the study. The patients' characteristics, including sex, age, recurrence, and death, were extracted from the medical records of the hospital.

The study was approved by the Ethics Committee of Mashhad University of Medical Sciences (code: [IR.MUMS.MEDICAL.REC.1400.689](#)), and ethical considerations were met throughout the study steps.

### 3.1. Statistical Analysis

The patients' characteristics were described using mean  $\pm$  standard deviation (SD) for age or frequency (percentage) for the categorical variables. The comparison of the pathologists' agreement was performed using the Kappa method. For comparing the agreement among the four pathologists, Fleiss Multirater Kappa was used for grade, stromal invasion, muscular invasion, and tumor heterogeneity separately. The chance-corrected observed agreement was calculated as:

$$\kappa = p \text{ observed} - p \text{ Expected} / 1 - p$$

A weighted average of the values for the individual categories was considered for the overall value of Kappa for more than two categories. The value of  $\kappa$  can range from -1.0 to +1.0 (0 indicates chance agreement and 1.0 indicates perfect agreement; negative values indicate systematic disagreement). As generally accepted,  $\kappa \geq 0.75$  was considered excellent agreement, 0.40 - 0.75 was fair to good agreement, and values  $< 0.40$  were poor agreement (20).

All statistical analyses were performed using IBM SPSS Statistics for Windows, version 23.0 (IBM Corp. 2014. Armonk, NY: IBM Corp.). P-values  $< 0.05$  were considered statistically significant.

#### 4. Results

A total of 132 samples were included in this study. These samples were from 21 women (15.9%) and 111 men (84.1%); the mean age of patients was  $66.56 \pm 11.31$  years (range 42 - 91 years). The patients were followed up for 9 - 46 months.

The analysis of Kappa showed significant agreement among the pathologists in all tumor characteristics ( $P < 0.001$ ; Table 1). There was fair to good agreement for the 2004/2016 WHO grading system ( $\kappa = 0.602$ ) and Cheng's grading system ( $\kappa = 0.574$ ), excellent agreement for stromal invasion ( $\kappa = 0.790$ ) and muscular invasion ( $\kappa = 0.884$ ), and poor agreement for tumor heterogeneity ( $\kappa = 0.360$ ).

We had no patients with grade I based on Cheng's grading system. Comparing the grading systems showed a  $\kappa = 0.602$  for both low and high grades, which was similar to the agreement for grade II in Cheng's grading system. However, the agreement was lower for grade III ( $\kappa = 0.439$ ) and higher for grade IV ( $\kappa = 0.690$ ) compared with the WHO 2004/2016 grading system.

In the next step, we compared the grading results of the four pathologists pairwise, and the results showed significant P values for all pairs in both the WHO grading system and Cheng's system ( $P < 0.001$ ; Table 2). As observed, most of the pairwise comparisons showed fair to good agreement between the pathologists, except for pathologist one vs. pathologist four, which showed poor agreement in Cheng's grading system.

#### 5. Discussion

The results of the present study indicated "fair to good" agreement for both grading systems, 2004/2016 WHO and Cheng's system, in general, and in subgroups for all grades, i.e., low grade and high grade in the 2004/2016 grading system and grades II, III, and IV in Cheng's grading system. This comparison has not been

performed previously; therefore, we cannot compare the results with similar studies, and we must compare our results with the few studies reporting the reproducibility of the 2004/2016 grading system compared with previous WHO systems. The agreement among the pathologists in the 2004/2016 system ranged from 39 - 74%, with kappa values of  $\kappa = 0.14 - 0.58$  (11),  $\kappa = 0.30 - 0.52$  (12), and  $\kappa = 0.35$  (15) in different studies. We have obtained higher Kappa values in the present study compared with the two studies mentioned above (11, 12, 15). The inclusion of a specific group of patients may be one of the reasons for such differences. Furthermore, this difference can be related to the specification of grade III in Cheng et al.'s system (8).

It should be noted that the tumor's grade is an important source of inter-observer variability; the overall inter-observer agreement of 87% ( $\kappa = 0.70$ ), reported by Mangrud et al. for the WHO 2004/2016 grading system, which evaluated patients with TaT1 carcinoma, decreased to 66% in high-grade tumors and increased to 100% agreement in low-grade tumors (16). The Kappa values reported in this study for the WHO 2004/2016 grading system were close to those reported in the present study, indicating "fair to good" inter-observer agreement. A review of 20 studies indicated a wide range of inter-observer agreement for the 2004/2016 classification (43 - 100%,  $\kappa = 0.17 - 0.70$ ) (13). This wide range refers to differences in the disease subgroup analyzed and the limitations of the studies (moderate to high risk of bias) (13).

In addition to the tumor's grading, we also investigated the pathologists' agreement in reporting stromal invasion, muscular invasion, and tumor heterogeneity; the results showed good agreement in all parameters except heterogeneity. Few studies have evaluated the pathologists' agreement for stromal invasion and none for tumor heterogeneity and muscular invasion; although it is well known these are essential for evaluating the depth of invasion and have a significant effect on disease progression and, thus, patients' prognosis (21-23). Tosoni and colleagues reported that the report of one pathologist about stromal invasion should not be considered for treatment choice and suggested its evaluation by another pathologist (21). In the study suggesting a new staging system for pT1 papillary bladder cancer (superficially invasive papillary urothelial cell carcinoma), stromal invasion was used for defining microinvasion, and the results showed 81% agreement between two pathologists (24). In another study on bladder cancer, full agreement was observed only in 44% of eight genitourinary pathologists, with greater

**Table 1.** The Agreement Between the Four Pathologists on Different Tumor Characteristics

Variables	Kappa	P-Value	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
<b>Stromal invasion</b>				
Overall agreement	0.790	< 0.001	0.720	0.859
No	0.790	< 0.001	0.720	-
Yes	0.790	< 0.001	0.720	-
<b>Muscular invasion</b>				
Overall agreement	0.884	< 0.001	0.813	0.954
No	0.884	< 0.001	0.813	-
Yes	0.884	< 0.001	0.813	-
<b>Grade (2016)</b>				
Overall agreement	0.602	< 0.001	0.532	0.672
Low grade	0.602	< 0.001	0.532	0.680
High grade	0.602	< 0.001	0.532	0.922
<b>Grade (new system)</b>				
Overall agreement	0.574	< 0.001	0.524	0.625
Grade II	0.602	< 0.001	0.532	-
Grade III	0.439	< 0.001	0.370	-
Grade IV	0.690	< 0.001	0.620	-
<b>Tumor heterogeneity</b>				
Overall agreement	0.360	< 0.001	0.291	0.430
No	0.360	< 0.001	0.291	0.853
Yes	0.360	< 0.001	0.291	-

**Table 2.** Pairwise Comparison of the Grading Results of the Four Pathologists

System	Pathologist 1 vs. 2	Pathologist 1 vs. 3	Pathologist 1 vs. 4	Pathologist 2 vs. 3	Pathologist 2 vs. 4	Pathologist 3 vs. 4
<b>WHO grading system</b>						
PUNLMP	0	0	0	0	0	0
Low grade	13	14	11	23	21	23
High grade	100	104	99	96	92	97
Kappa	0.506	0.612	0.415	0.716	0.595	0.679
<b>Cheng's grading system</b>						
Grade I	0	0	0	0	0	0
Grade II	12	14	11	22	20	23
Grade III	39	32	27	31	31	28
Grade IV	38	45	39	45	44	54
Kappa	0.484	0.511	0.345	0.607	0.572	0.679

discordance between aggressive and conservative pathologists (25). Others have also identified lamina propria invasion as one of the parameters with poor agreement between experienced pathologists (26), while we found good agreement between pathologists in this regard. Tumor heterogeneity, which refers to the divergent differentiation of this type of cancer, is identified as another cause of observer variability, which can also result in uncertainty of treatment choice (11, 27, 28). Therefore, further studies are required to report the

pathologists' agreement on these parameters to be comparable to the results of the present study.

The characteristics of the pathologists may also affect the results of our study. In our study, pairwise comparisons showed no difference in Kappa values, with fair to good agreement between the pathologists in both grading systems, except in one comparison. Some have suggested that the level of experience is an important factor in poor inter-observer agreement for grading urothelial carcinoma on urine cytology,

reporting more accuracy in reports from senior pathologists with more than seven years of experience or those with specialized training (29). Others reported no difference in Kappa values for pathologists with more than ten years versus less than ten years of experience in low-grade urothelial carcinoma of instrumented urinary tract cytology specimens (30). Possibly, experience with that specific grading system is more important than the overall years of experience of the pathologists (31). We cannot comment on this based on our study because we did not have any pathologists with a low level of experience. We also evaluated whether the medical center where the pathologists work can influence their agreement, but the results rejected such an effect, and all Kappa values fell within one category (fair to moderate). Further studies are required to determine whether the pathologists' characteristics can influence the accuracy of the results reported for grading UCB.

Regarding the patients' characteristics, most patients in our study were men, which is consistent with previous reports. This sex difference is attributed to the exposure to risk factors, such as cigarette smoking and occupational hazards in men, as well as the role of sex hormones (32), with different male-to-female ratios reported (33). Additionally, the mean age of our study population was close to that reported in the USA (34) and Canada (35). It has been suggested that patients younger than 40 years have smaller tumors, lower-grade cancers, and better overall survival (36, 37), which may explain why we had no grade I pathology (in the new grading system) among our samples and the poor prognosis observed.

The main limitation of this study was the retrospective nature of data collection, which resulted in a large amount of missing data on some variables, such as clinical outcomes (death and recurrence); however, these variables were not the main objective of our study and thus not critical. Another consequence of retrospective evaluation was that we could not differentiate between diagnoses based on biopsy versus surgical specimens. There were also several variables not available in the medical records and not evaluated, such as underlying diseases, whose confounding effects may influence the results of our study. Additionally, we did not collect any data about the treatment strategies employed; therefore, the results of survival and recurrence could not be interpreted completely. For the same reason, we did not comment on the impact of these grading systems on the therapeutic strategies selected by the oncologists. Finally, we selected patients from one medical center; therefore, the results cannot

be generalized to all patients with this condition, and further research is required to generalize the results to a wider population.

### 5.1. Conclusions

The fair to good agreement among pathologists in both grading systems for UCB, namely the WHO 2004/2016 and Cheng's grading system, and the close Kappa values demonstrated the reproducibility of these two grading systems and rejected differences in the grading systems as the cause of discrepancies among pathologists. Therefore, the choice of grading system used for the pathological report of a UCB specimen should not be based on the reproducibility of the grading system. Other factors may be significantly different between the two grading systems, the determination of which requires further studies.

### Footnotes

**Authors' Contribution:** S.T.J.: Acquisition of data, analysis and interpretation of data, critical revision of the manuscript for important intellectual content. M.S. Study concept and design, administrative, technical, and material support, drafting of the manuscript, critical revision of the manuscript for important intellectual content. M.R.K.: Acquisition of data, analysis and interpretation of data, critical revision of the manuscript for important intellectual content. M.F.: Acquisition of data, analysis and interpretation of data, critical revision of the manuscript for important intellectual content. A.J.: Acquisition of data, analysis and interpretation of data, critical revision of the manuscript for important intellectual content. N.M.: Statistical analysis, critical revision of the manuscript for important intellectual content. All the authors reviewed and approved the final version of the manuscript.

**Conflict of Interests Statement:** The authors declare that there is no conflict of interest.

**Data Availability:** The dataset presented in the study is available on request from the corresponding author during submission or after publication.

**Ethical Approval:** The study was approved by the Ethics Committee of Mashhad University of Medical Sciences (code: [IR.MUMS.MEDICAL.REC.1400.689](#)), and ethical considerations were met throughout the study steps.

**Funding/Support:** This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

**Informed Consent:** Not declared by the authors.

## References

- Miyazaki J, Nishiyama H. Epidemiology of urothelial carcinoma. *Int J Urol.* 2017;**24**(10):730-4. [PubMed ID: 28543959]. <https://doi.org/10.1111/iju.13376>.
- Froemming A, Potretzke T, Takahashi N, Kim B. Upper tract urothelial cancer. *Eur J Radiol.* 2018;**98**:50-60. [PubMed ID: 29279170]. <https://doi.org/10.1016/j.ejrad.2017.10.021>.
- Farina MS, Lundgren KT, Bellmunt J. Immunotherapy in Urothelial Cancer: Recent Results and Future Perspectives. *Drugs.* 2017;**77**(10):1077-89. [PubMed ID: 28493171]. <https://doi.org/10.1007/s40265-017-0748-7>.
- Jn E. Pathology and genetics of tumours of the urinary system and male genital organs. *WHO classification of tumors.* 2004.
- Comperat EM, Burger M, Gontero P, Mostafid AH, Palou J, Roupret M, et al. Grading of Urothelial Carcinoma and The New "World Health Organisation Classification of Tumours of the Urinary System and Male Genital Organs 2016". *Eur Urol Focus.* 2019;**5**(3):457-66. [PubMed ID: 29366854]. <https://doi.org/10.1016/j.euf.2018.01.003>.
- Moch H, Cubilla AL, Humphrey PA, Reuter VE, Ulbright TM. The 2016 WHO Classification of Tumours of the Urinary System and Male Genital Organs-Part A: Renal, Penile, and Testicular Tumours. *Eur Urol.* 2016;**70**(1):93-105. [PubMed ID: 26935559]. <https://doi.org/10.1016/j.eururo.2016.02.029>.
- Humphrey PA, Moch H, Cubilla AL, Ulbright TM, Reuter VE. The 2016 WHO Classification of Tumours of the Urinary System and Male Genital Organs-Part B: Prostate and Bladder Tumours. *Eur Urol.* 2016;**70**(1):106-19. [PubMed ID: 26996659]. <https://doi.org/10.1016/j.eururo.2016.02.028>.
- Cheng L, MacLennan GT, Lopez-Beltran A. Histologic grading of urothelial carcinoma: a reappraisal. *Hum Pathol.* 2012;**43**(12):2097-108. [PubMed ID: 22542126]. <https://doi.org/10.1016/j.humpath.2012.01.008>.
- MacLennan GT, Kirkali Z, Cheng L. Histologic grading of noninvasive papillary urothelial neoplasms. *Eur Urol.* 2007;**51**(4):889-97. discussion 897-8. [PubMed ID: 17095142]. <https://doi.org/10.1016/j.eururo.2006.10.037>.
- Engers R. Reproducibility and reliability of tumor grading in urological neoplasms. *World J Urol.* 2007;**25**(6):595-605. [PubMed ID: 17828603]. <https://doi.org/10.1007/s00345-007-0209-0>.
- van Rhijn BW, van Leenders GJ, Ooms BC, Kirkels WJ, Zlotta AR, Boeve ER, et al. The pathologist's mean grade is constant and individualizes the prognostic value of bladder cancer grading. *Eur Urol.* 2010;**57**(6):1052-7. [PubMed ID: 19765886]. <https://doi.org/10.1016/j.eururo.2009.09.022>.
- May M, Brookman-Amissh S, Roigas J, Hartmann A, Storkel S, Kristiansen G, et al. Prognostic accuracy of individual uropathologists in noninvasive urinary bladder carcinoma: a multicentre study comparing the 1973 and 2004 World Health Organization classifications. *Eur Urol.* 2010;**57**(5):850-8. [PubMed ID: 19346063]. <https://doi.org/10.1016/j.eururo.2009.03.052>.
- Soukup V, Capoun O, Cohen D, Hernandez V, Babjuk M, Burger M, et al. Prognostic Performance and Reproducibility of the 1973 and 2004/2016 World Health Organization Grading Classification Systems in Non-muscle-invasive Bladder Cancer: A European Association of Urology Non-muscle Invasive Bladder Cancer Guidelines Panel Systematic Review. *Eur Urol.* 2017;**72**(5):801-13. [PubMed ID: 28457661]. <https://doi.org/10.1016/j.eururo.2017.04.015>.
- Jones TD, Cheng L. Papillary urothelial neoplasm of low malignant potential: evolving terminology and concepts. *J Urol.* 2006;**175**(6):1995-2003. [PubMed ID: 16697785]. [https://doi.org/10.1016/S0022-5347\(06\)00267-9](https://doi.org/10.1016/S0022-5347(06)00267-9).
- Bosschieter J, Hentschel A, Savci-Heijink CD, Patrick van der Voorn J, Rozendaal L, Vis AN, et al. Reproducibility and Prognostic Performance of the 1973 and 2004 World Health Organization Classifications for Grade in Non-muscle-invasive Bladder Cancer: A Multicenter Study in 328 Bladder Tumors. *Clin Genitourin Cancer.* 2018;**16**(5):e985-92. [PubMed ID: 29884516]. <https://doi.org/10.1016/j.clgc.2018.05.002>.
- Mangrud OM, Waalen R, Gudlaugsson E, Dalen I, Tasdemir I, Janssen EA, et al. Reproducibility and prognostic value of WHO1973 and WHO2004 grading systems in TaT1 urothelial carcinoma of the urinary bladder. *PLoS One.* 2014;**9**(1): e83192. [PubMed ID: 24409280]. [PubMed Central ID: PMC3883638]. <https://doi.org/10.1371/journal.pone.0083192>.
- Bostwick DG, Cheng L. *Urologic surgical pathology.* Elsevier Health Sciences; 2008.
- Cheng L, Neumann RM, Nehra A, Spotts BE, Weaver AL, Bostwick DG. Cancer heterogeneity and its biologic implications in the grading of urothelial carcinoma. *Cancer.* 2000;**88**(7):1663-70. [PubMed ID: 10738225]. [https://doi.org/10.1002/\(sici\)1097-0142\(20000401\)88:73.0.co;2-8](https://doi.org/10.1002/(sici)1097-0142(20000401)88:73.0.co;2-8).
- Culpan M, Keser F, Iplikci A, Kir G, Atis G, Yildirim A. The Clinical Impact of Tumor Grade Heterogeneity in Non-muscle-invasive Urothelial Carcinoma of the Bladder. *Medeni Med J.* 2021;**36**(4):310-7. [PubMed ID: 34937341]. [PubMed Central ID: PMC8694165]. <https://doi.org/10.4274/MMJ.galenos.2021.48447>.
- Mamoon N, Iqbal MA, Jamal S, Luqman M. Urothelial neoplasia of the urinary bladder—comparison of interobserver variability for WHO Classification 1972 with WHO/ISUP Consensus Classification 1998. *J Ayub Med Coll Abbottabad.* 2006;**18**(2):4-8. [PubMed ID: 16977804].
- Tosoni I, Wagner U, Sauter G, Egloff M, Knonagel H, Alund G, et al. Clinical significance of interobserver differences in the staging and grading of superficial bladder cancer. *BJU Int.* 2000;**85**(1):48-53. [PubMed ID: 10619945]. <https://doi.org/10.1046/j.1464-410x.2000.00356.x>.
- Comperat E, Varinot J, Moroch J, Eymerit-Morin C, Brimo F. A practical guide to bladder cancer pathology. *Nat Rev Urol.* 2018;**15**(3):143-54. [PubMed ID: 29384523]. <https://doi.org/10.1038/nrurol.2018.2>.
- Hansel DE, Amin MB, Comperat E, Cote RJ, Knuchel R, Montironi R, et al. A contemporary update on pathology standards for bladder cancer: transurethral resection and radical cystectomy specimens. *Eur Urol.* 2013;**63**(2):321-32. [PubMed ID: 23088996]. <https://doi.org/10.1016/j.eururo.2012.10.008>.
- van der Aa MN, van Leenders GJ, Steyerberg EW, van Rhijn BW, Jobsis AC, Zwarthoff EC, et al. A new system for substaging pT1 papillary bladder cancer: a prognostic evaluation. *Hum Pathol.* 2005;**36**(9):981-6. [PubMed ID: 16153461]. <https://doi.org/10.1016/j.humpath.2005.06.017>.
- Comperat E, Egevad L, Lopez-Beltran A, Camparo P, Algaba F, Amin M, et al. An interobserver reproducibility study on invasiveness of bladder cancer using virtual microscopy and heatmaps. *Histopathology.* 2013;**63**(6):756-66. [PubMed ID: 24102813]. <https://doi.org/10.1111/his.12214>.
- Lopez-Beltran A, Cheng L. Stage T1 bladder cancer: diagnostic criteria and pitfalls. *Pathology.* 2021;**53**(1):67-85. [PubMed ID: 33153725]. <https://doi.org/10.1016/j.pathol.2020.09.014>.
- Comperat E, Oszwald A, Wasinger G, Hansel DE, Montironi R, van der Kwast T, et al. Updated pathology reporting standards for bladder cancer: biopsies, transurethral resections and radical cystectomies. *World J Urol.* 2022;**40**(4):915-27. [PubMed ID: 34554298]. [PubMed Central ID: PMC8994708]. <https://doi.org/10.1007/s00345-021-03831-1>.

28. Cheng L, Neumann RM, Weaver AL, Cheville JC, Leibovich BC, Ramnani DM, et al. Grading and staging of bladder carcinoma in transurethral resection specimens. Correlation with 105 matched cystectomy specimens. *Am J Clin Pathol.* 2000;**113**(2):275-9. [PubMed ID: 10664630]. <https://doi.org/10.1309/94B6-8VFB-MN9J-1NF5>.
29. Reid MD, Osunkoya AO, Siddiqui MT, Looney SW. Accuracy of grading of urothelial carcinoma on urine cytology: an analysis of interobserver and intraobserver agreement. *Int J Clin Exp Pathol.* 2012;**5**(9):882-91. [PubMed ID: 23119105]. [PubMed Central ID: PMC3484494].
30. McCroskey Z, Pambuccian SE, Kleitherms S, Antic T, Cohen MB, Barkan GA, et al. Accuracy and interobserver variability of the cytologic diagnosis of low-grade urothelial carcinoma in instrumented urinary tract cytology specimens. *Am J Clin Pathol.* 2015;**144**(6):902-8. [PubMed ID: 26572997]. <https://doi.org/10.1309/AJCPEIO9YKMRSQKG>.
31. Bulten W, Pinckaers H, van Boven H, Vink R, de Bel T, van Ginneken B, et al. Automated deep-learning system for Gleason grading of prostate cancer using biopsies: a diagnostic study. *Lancet Oncol.* 2020;**21**(2):233-41. [PubMed ID: 31926805]. [https://doi.org/10.1016/S1470-2045\(19\)30739-9](https://doi.org/10.1016/S1470-2045(19)30739-9).
32. Goto T, Miyamoto H. The Role of Estrogen Receptors in Urothelial Cancer. *Front Endocrinol (Lausanne).* 2021;**12**:643870. [PubMed ID: 33796076]. [PubMed Central ID: PMC8008958]. <https://doi.org/10.3389/fendo.2021.643870>.
33. Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin DM, Pineros M, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *Int J Cancer.* 2019;**144**(8):1941-53. [PubMed ID: 30350310]. <https://doi.org/10.1002/ijc.31937>.
34. Raman JD, Messer J, Sielatycki JA, Hollenbeak CS. Incidence and survival of patients with carcinoma of the ureter and renal pelvis in the USA, 1973-2005. *BJU Int.* 2011;**107**(7):1059-64. [PubMed ID: 20825397]. <https://doi.org/10.1111/j.1464-410X.2010.09675.x>.
35. Abouassaly R, Alibhai SM, Shah N, Timilshina N, Fleshner N, Finelli A. Troubling outcomes from population-level analysis of surgery for upper tract urothelial carcinoma. *Urology.* 2010;**76**(4):895-901. [PubMed ID: 20646743]. <https://doi.org/10.1016/j.urology.2010.04.020>.
36. Telli O, Sarici H, Ozgur BC, Doluoglu OG, Sunay MM, Bozkurt S, et al. Urothelial cancer of bladder in young versus older adults: clinical and pathological characteristics and outcomes. *Kaohsiung J Med Sci.* 2014;**30**(9):466-70. [PubMed ID: 25224770]. <https://doi.org/10.1016/j.kjms.2014.02.017>.
37. Woodford R, Ranasinghe W, Aw HC, Sengupta S, Persad R. Trends in incidence and survival for upper tract urothelial cancer (UTUC) in the state of Victoria–Australia. *BJU Int.* 2016;**117** Suppl 4:45-9. [PubMed ID: 26486693]. <https://doi.org/10.1111/bju.13232>.