



Comparison between 1973 and 2004/2016 World Health Organization grading in upper tract urothelial carcinoma treated with radical nephroureterectomy

Claudia Collà Ruvolo^{1,2} · Christoph Würnschimmel^{1,3,4} · Mike Wenzel^{1,5} · Luigi Nocera^{1,3} · Giuseppe Celentano² · Francesco Mangiapia² · Zhe Tian¹ · Shahrokh F. Shariat^{6,7,8,9,10,11} · Fred Saad¹ · Felix H. C. Chun⁵ · Alberto Briganti³ · Nicola Longo² · Vincenzo Mirone² · Pierre I. Karakiewicz¹

Received: 23 January 2021 / Accepted: 21 May 2021 / Published online: 6 June 2021
© The Author(s) 2021

Abstract

Aims The European Association of Urology guideline for upper tract urothelial carcinoma (UTUC) relies on two grading system: 1973 World Health Organization (WHO) and 2004/2016 WHO. No consensus has been made which classification should supersede the other and both are recommended in clinical practice. We hypothesized that one may be superior to the other.

Methods Newly diagnosed non-metastatic UTUC patients treated with radical nephroureterectomy were abstracted from the Surveillance, Epidemiology, and End Results database (2010–2016). Kaplan–Meier plots and multivariable Cox regression models (CRMs) tested cancer-specific mortality (CSM), according to 1973 WHO (G_1 vs. G_2 vs. G_3) or to 2004/2016 WHO (low-grade vs. high-grade) grading systems. Haegerty’s C-index quantified accuracy.

Results Of 4271 patients, according to 1973 WHO grading system, 134 (3.1%) were G_1 , 436 (10.2%) were G_2 and 3701 (86.7%) were G_3 ; while according to 2004/2016 WHO grading system, 508 (11.9%) were low grade vs 3763 (88.1%) high grade. In multivariable CRMs, high grade predicted higher CSM (Hazard ratio: 1.70, $p < 0.001$). Conversely, neither G_2 ($p = 0.8$) nor G_3 ($p = 0.1$) were independent predictors of worse survival. The multivariable models without consideration of either grading system were 74% accurate in predicting 5-year CSM. Accuracy increased to 76% after either addition of the 1973 WHO or 2004/2016 WHO grade.

Conclusions From a statistical standpoint, either 1973 WHO or 2004/2016 WHO grading system improves the accuracy of CSM prediction to the same extent. In consequence, other considerations such as intra- and interobserver variability may represent additional metrics to consider in deciding which grading system is better.

Keywords WHO · Grade · UTUC · SEER · Pathology

✉ Claudia Collà Ruvolo
c.collaruvolo@gmail.com

¹ Cancer Prognostics and Health Outcomes Unit, Division of Urology, University of Montréal Health Center, Montreal, QC, Canada

² Urology Unit, Department of Neurosciences, Reproductive Sciences and Odontostomatology, University of Naples Federico II, Naples, Italy

³ Department of Urology and Division of Experimental Oncology, URI, Urological Research Institute, IBCAS San Raffaele Scientific Institute, Milan, Italy

⁴ Martini-Klinik Prostate Cancer Center, University Hospital Hamburg-Eppendorf, Hamburg, Germany

⁵ Department of Urology, University Hospital Frankfurt, Frankfurt am Main, Germany

⁶ Department of Urology, Comprehensive Cancer Center, Medical University of Vienna, Vienna, Austria

⁷ Departments of Urology, Weill Cornell Medical College, New York, NY, USA

⁸ Department of Urology, University of Texas Southwestern, Dallas, TX, USA

⁹ Department of Urology, Second Faculty of Medicine, Charles University, Prag, Czech Republic

¹⁰ Institute for Urology and Reproductive Health, I.M. Sechenov First Moscow State Medical University, Moscow, Russia

¹¹ Division of Urology, Department of Special Surgery, Jordan University Hospital, The University of Jordan, Amman, Jordan

Introduction

Upper tract urothelial carcinoma (UTUC) is a rare and aggressive malignancy, with an estimated annual incidence in Western Countries of almost two cases per 100,000 inhabitants [1] and with non-organ confined stage in two-third of newly diagnosed patients [2–5]. After stage, tumor grade is the most important predictor of cancer-specific mortality (CSM) in UTUC patients [6–10]. The most recent European Association of Urology (EAU) UTUC guideline relies and recommends the use of two different grading systems. These consist of the 1973 World Health Organization (WHO) and the 2004/2016 WHO classification. Specifically, the 1973 WHO grading system [11] is based on three tiers. Grade 1 applies to tumors with least degree of cellular anaplasia. Grade 3 applies to tumors with most severe degrees of cellular anaplasia. Finally, grade 2 lies in between. Conversely, the 2004/2016 WHO grading system [12, 13] is based on two tiers. It relies on more detailed histological criteria. Low-grade carcinoma applies to tumors with predominantly ordered cell organization with mainly round–oval nuclear shape and mild nuclear chromatin variation. High grade applies to tumors with predominantly disordered cell organization with loss of polarity, moderate to marked nuclear pleomorphism and mainly hyperchromasia [14]. Since there is no consensus on which of the two grading systems should be used in everyday clinical practice [12, 15] and since both are recommended [2], we hypothesized that one may be better. To test this hypothesis, we examined the ability of either the 1973 or the 2004/2016 WHO grading system in predicting CSM, in a contemporary cohort of non-metastatic UTUC patients treated with radical nephroureterectomy (RNU), identified within a large-scale database, namely the Surveillance, Epidemiology and End Results, from 2010 to 2016.

Materials and methods

Study population

The 2019-release SEER-18 registry database covers 34.6% of the United States population [16]. Within SEER-18 database (2010–2016), we identified patients aged ≥ 18 year, diagnosed with primary histologically confirmed urothelial carcinoma of renal pelvis or ureter [International Classification of Disease for Oncology (ICD-O-3) site code C65.9 and C.66.9] and treated with RNU. Autopsy and death certificate only cases, with other histology than urothelial ($n = 142$), distant metastases ($n = 323$), unknown T-stage ($n = 43$) and unknown grade ($n = 813$) were excluded. These inclusion criteria yielded 4271 patients.

Variables definition

Tumor grade was defined according to both the 1973 WHO grading system [grade 1 (G_1) vs. grade 2 (G_2) vs. grade 3 (G_3)] and the 2004/2016 WHO grading system (low grade vs. high grade). Covariables consisted of age, sex, primary site (renal pelvis, ureter), T-stage (T_1 vs. T_2 vs. T_3 vs. T_4), N-stage (N_0 vs. N_+ vs. N_x) and chemotherapy administration (yes vs. no/unknown). CSM was defined as deaths related to UTUC, according to SEER mortality code [17] and represented the endpoint of interest.

Statistical analyses

Kaplan–Meier plots and multivariable Cox regression models predicting CSM were fitted. These models relied on T-stage, N-stage, chemotherapy administration and primary site, without including grade. Subsequently, the models were refitted with all previously included variables in addition to the 1973 WHO grading system. Finally, the models were refitted again, this time, with the 2004/2016 WHO grading system. Within Cox models, independent predictor status of WHO grading system was tested. Sensitivity analyses testing the effect of grade (1973 and 2004/2016 WHO grading systems) on CSM were performed in UTUC patients with T_1 stage and in UTUC patients with T_2 or lower stage. Finally, the effect of 2004/2016 WHO grading system on CSM was tested in UTUC patients with G_2 grade, according to the 1973 WHO grading system. Subsequently, accuracy of 5-year CSM predictions was quantified based on multivariable models without consideration of WHO grading system, as well as with consideration of either the 1973 or the 2004/2016 WHO grading system. Haegerty's C-index quantified accuracy. All statistical tests were two sided, with a level of significance set at $p < 0.05$. Statistical analyses were performed using the R software environment for statistical computing and graphics, version 4.0.0 (available at: <http://www.rproject>).

Results

Descriptive characteristics

From 2010 to 2016, 4271 cases of UTUC treated with RNU were identified (Table 1). Of those, according to 1973 WHO grading system, 134 (3.1%) were G_1 , 436 (10.2%) were G_2 and 3701 (86.7%) were G_3 ; while according to 2004/2016 WHO grading system, 508 (11.9%) were low grade vs 3763 (88.1%) high grade. The median age was 73 years (Interquartile range: 65–80). Most patients were male ($n = 2575$, 60.3%), with renal pelvis urothelial carcinoma ($n = 2906$,

Table 1 Baseline characteristics of 4271 upper tract urothelial carcinoma patients treated with radical nephroureterectomy, identified within Surveillance, Epidemiology and End Results database, between 2010 and 2016

Overall, <i>n</i> (%)	4271 (100)
1973 WHO grading system, <i>n</i> (%)	
G ₁	134 (3.1)
G ₂	436 (10.2)
G ₃	3701 (86.7)
2004/2016 WHO grading system, <i>n</i> (%)	
Low grade	508 (11.9)
High grade	3763 (88.1)
Age	
Median	73
IQR	65–80
Sex, <i>n</i> (%)	
Female	1696 (39.7)
Male	2575 (60.3)
Follow-up	
Median	22
IQR	10–43
Primary site, <i>n</i> (%)	
Renal pelvis	2906 (68.0)
Ureter	1365 (32.0)
T-stage, <i>n</i> (%)	
T ₁	1306 (30.6)
T ₂	747 (17.5)
T ₃	1867 (43.7)
T ₄	351 (8.2)
N-stage, <i>n</i> (%)	
N ₀	3711 (86.9)
N ₊	476 (11.1)
N _x	84 (2.0)
Chemotherapy, <i>n</i> (%)	
No/unknown	3374 (79.0)
Yes	897 (21.0)

IQR interquartile range, WHO World Health Organization

68.0%) and harbored T₃ stage at RNU (*n* = 1867, 43.7%). Finally, 897 (21.0%) patients received chemotherapy. Of all G₁ patients (*n* = 134), 119 (88.8%) and 15 (11.2%) were low grade and high grade, respectively. Of all G₂ patients (*n* = 436), 358 (82.1%) and 78 (17.9) were low grade and high grade, respectively. Of all G₃ patients (*n* = 3701), 31 (0.8%) and 3670 (99.2%) were low grade and high grade, respectively (Fig. 1).

Survival analyses and accuracy in predicting CSM

In overall population, according to 1973 WHO grading system (Fig. 2A), 5-year CSM rates were 10.2%, 14.6% and 30.5% for G₁, G₂ and G₃ UTUC grade, respectively.

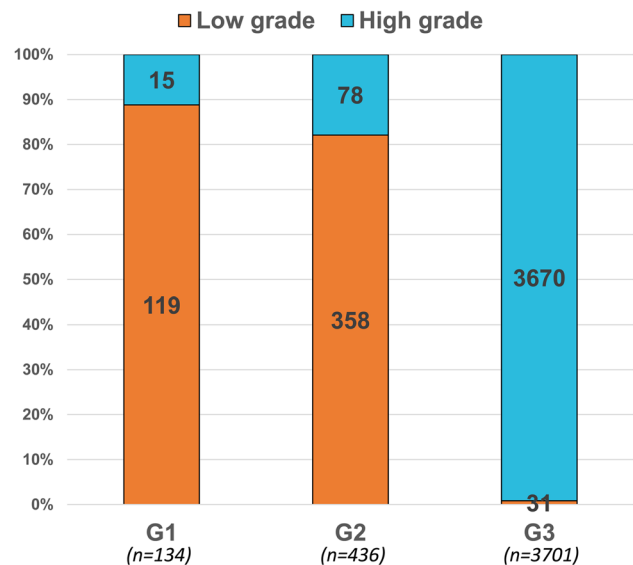


Fig. 1 Stacked barplot depicting the rates of tumor grade according to the 2004/2016 WHO grading system (low grade vs high grade) in 134, 436 and 3701 G₁, G₂ and G₃ non-metastatic upper tract urothelial carcinoma patients treated with radical nephroureterectomy, according to the 1973 WHO grading system, respectively

In multivariable Cox regression models focusing on CSM (Table 2), relative to G₁, neither G₂ [Hazard ratio (HR) 1.07, *p* = 0.8] or G₃ (HR 1.65, *p* = 0.1) represented independent predictors. When sensitivity analyses were performed (Supplementary Table 1), the results were confirmed in the multivariable Cox regression models focusing on CSM in patients with T₁ stage (relative to G₁, G₂: HR 1.00, *p* = 1.0 and G₃: HR 1.82, *p* = 0.2) and T₂ or lower stage (G₂ HR: 0.99 *p* = 0.9, G₃ HR 1.38, *p* = 0.4, relative to G₁). The accuracy of the multivariable model (Table 4) that included 1973 WHO grading system was 76%. Conversely, the accuracy of the multivariable model without consideration of 1973 WHO grading system was 74%.

In overall population, according to 2004/2016 WHO grading system (Fig. 2B), 5-year CSM rates were 13.4% and 30.2% for low grade and high grade, respectively. In multivariable Cox regression models focusing on CSM (Table 3), relative to low grade, high grade (HR 1.70, *p* < 0.001) achieved independent predictor status. When sensitivity analyses were performed (Supplementary Table 1), the results were confirmed in the multivariable Cox regression models focusing on CSM in patients with T₁ stage (relative to low grade, high grade: HR 1.76, *p* = 0.04), T₂ or lower stage (relative to low grade, high grade: HR 1.65, *p* = 0.02) and G₂ grade (relative to low grade, high grade: HR 2.19, *p* = 0.02). The accuracy of the multivariable model (Table 4) that included 2004/2016 WHO grading system was 76%. Conversely, the accuracy of the multivariable model without consideration of 2004/2016 WHO grading system was 74%.

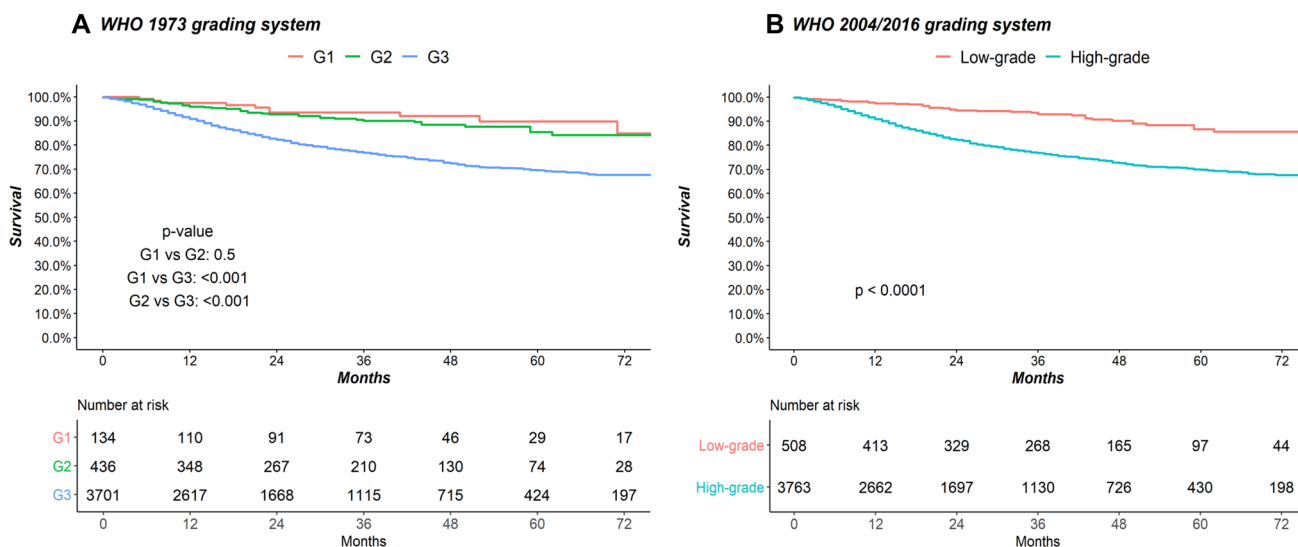


Fig. 2 Kaplan–Meier plots depicting cancer-specific mortality (CSM) in 4271 non-metastatic upper tract urothelial carcinoma patients treated with radical nephroureterectomy, identified within Surveil-

lance, Epidemiology and End Results (2010–2016), according to the **A** 1973 World Health Organization (WHO) grading system and to the **B** 2004/2016 WHO grading system

Table 2 Multivariable Cox regression models predicting cancer-specific mortality (CSM) in 4271 upper tract urothelial carcinoma patients identified within Surveillance, Epidemiology and End Results database (2010–2016), where pathological grade was defined according to the three-tier 1973 World Health Organization (WHO) grading system

	CSM	
	HR (95% CI)	p-value
WHO 1973 grading system, relative to G1		
G ₂	1.07 (0.55–2.08)	0.8
G ₃	1.65 (0.90–3.01)	0.1
T-stage, relative to T1		
T ₂	1.85 (1.39–2.48)	<0.001
T ₃	3.64 (2.88–4.60)	<0.001
T ₄	10.69 (8.06–14.17)	<0.001
N-stage, relative to N ₀		
N ₊	1.82 (1.49–2.22)	<0.001
N _x	0.79 (0.39–1.59)	0.5
Chemotherapy administration, relative to no/unknown		
Yes	0.80 (0.67–0.96)	0.01
Primary site, relative to renal pelvis		
Ureter	1.19 (1.01–1.40)	0.03

Discussion

To date, the EAU UTUC guideline relies and recommends the use of two different grade classification system: 1973 WHO and 2004/2016 WHO grading system. Which system should be used in everyday clinical practice is still under debate. We hypothesized that one may be better. To test this hypothesis, we examined the ability of either the 1973 or the 2004/2016 WHO grading system in predicting CSM, in a cohort of non-metastatic UTUC patients treated with RNU. Our analyses showed several noteworthy observations.

First, of all RNU patients examined in the current study ($n=4271$), approximately 90% harbored the highest grade level, regardless of which grading system was used. Specifically, 86.7% harbored G₃ according to 1973 WHO grading system and 88.1% harbored high grade according to 2004/2016 WHO grading system. These elevated rates of high-grade UTUC may be explained by the nature of the study population. Specifically, all patients harbored stage T₁ or higher [18]. Moreover, all patients were treated with RNU. In consequence, a selection bias towards higher grade was operational, relative to studies that also included non-invasive (stages T_a and T_{is}) UTUC patients treated with less

Table 3 Multivariable Cox regression models predicting cancer-specific mortality (CSM) in 4271 upper tract urothelial carcinoma patients identified within Surveillance, Epidemiology and End Results database (2010–2016), where pathological grade was defined according to the two-tier 2004/2016 World Health Organization (WHO) grading system

	CSM	
	HR (95% CI)	<i>p</i> -value
WHO 2004/2006 grading system, relative to low grade		
High grade	1.70 (1.23–2.35)	0.001
T-stage, relative to T ₁		
T ₂	1.84 (1.38–2.46)	<0.001
T ₃	3.61 (2.86–4.56)	<0.001
T ₄	10.51 (7.92–13.94)	<0.001
N-stage, relative to N ₀		
N ₊	1.82 (1.49–2.23)	<0.001
N _x	0.79 (0.39–1.60)	0.5
Chemotherapy administration, relative to no/unknown		
Yes	0.80 (0.67–0.96)	0.01
Primary site, relative to renal pelvis		
Ureter	1.19 (1.01–1.40)	0.04

Table 4 Accuracy in cancer-specific mortality prediction at 5 years after treatment, in 4217 upper tract urothelial carcinoma patients treated with radical nephroureterectomy, identified within Surveillance Epidemiology and End Results database (2010–2016), based on

multivariable Cox models: (1) without grade consideration, (2) considering the three-tier 1973 WHO grading system and (3) considering the two-tier 2004/2016 WHO grading system

	Heagerty's C-index
(1) Model based on primary site, T-stage, N-stage and chemotherapy administration	0.74
(2) Model based on primary site, T-stage, N-stage, chemotherapy administration with the three-tier <u>WHO 1973</u> grading classification system	0.76
(3) Model based on primary site, T-stage, N-stage, chemotherapy administration with the two-tier <u>WHO 2004/2016</u> grading classification system	0.76

WHO World Health Organization, *C-index* concordance index

definitive modalities than RNU [19–22]. However, even in those studies, the rate of non-invasive UTUC represented a marginal fraction of the overall population and the vast majority also harbored high-grade disease. For example, Singla et al. [21] examined 753 UTUC patients treated with RNU or distal ureterectomy, between 1998 and 2015. Of those, 78.8% harbored T₁ or higher stages and 89.2% harbored high-grade UTUC. Moreover, Roupret et al. [22] recorded T₁ or higher stages in 66 (68.0%) patients and high grade in 50 (51.5%) patients, within 97 UTUC patients, despite ureteroscopy or percutaneous endoscopy treatment.

Second, the current analyses demonstrated marginal discrimination between G₁ and G₂, with respect to CSM. Within the three-tier grading system, independent predictor status of G₂ and G₃, relative to G₁, could not be established. These results were confirmed in RNU patients with T₁ or T₂ or lower stages. The combination of these observations suggested limited discrimination ability of the three-tier grading system. Nonetheless, the addition of the 1973 WHO grading system resulted in a 2% accuracy gain, relative to

multivariable models without consideration of the three-tier grading system. However, a 2% gain may be considered marginal. Specifically, this figure implies that within a cohort of 1000 individuals, the use of the three-tier grading system would improve CSM prediction in 20 patients. This gain is important in large-scale prospective trials or in large-scale epidemiological analyses. However, a 2% gain in predictive accuracy may not be clinically meaningful in everyday clinical practice.

In the second part of the analyses, we focused on the two-tier WHO grading system. Here, we validated the independent predictor status of high grade relative to low grade. Specifically, high-grade UTUC had 1.70-fold, 1.76-fold, 1.65-fold, and 2.19-fold higher risk of CSM, relative to low-grade UTUC in overall population, in T₁, T₂ or lower and G₂ patients, respectively. Finally, we also recorded a 2% accuracy gain, when the 2004/2016 WHO grading system was added to multivariable model, where grade was previously not considered. In consequence, based on accuracy, the added benefit of the 2004/2016 WHO grading system

was exactly the same as for the 1973 WHO grading system. However, the discrimination of CSM rates appeared more practical with the two-tier grading system, where high-grade patients exhibited a nearly twofold higher CSM rate and reached independent predictor status. In consequence, it appears that based on statistical criteria used in the current analyses, the two-tier grading system benefits of a slight advantage over its three-tier counterpart.

Additional consideration may be required to decide which grading system should be included in everyday clinical practice and which may be abandoned. Several investigators compared intra- and interobserver variability of the two- vs three-tier grading system in bladder cancer [12, 23–29]. Unfortunately, such analyses did not focus on UTUC. However, based on methodological considerations, a system that relies on two tiers is invariably more likely to result in a lower intra- and interobserver variability than a system with more than two levels. This notion rests on the effect of chance. In consequence, based on similar predictive accuracy, superiority of discrimination in univariable and multivariable models, and on methodological consideration of intra- and interobserver variability, it appears that the two-tier grading system might represent a better alternative. However, specific expert intra- and interobserver variability testing in UTUC patients should ideally complement the findings of our study.

To the best of our knowledge, we are the first to examine the ability of either 1973 or 2004/2016 WHO grading classification in predicting CSM, in UTUC patients identified within a large-scale population-based database. Only one group of investigators [30] examined grade assignment differences according to 1973 vs. 2004/2016 grading system in a smaller cohort ($n=458$) of UTUC patients treated with RNU, at a single Chinese institution, between 2008 and 2013. Unfortunately, the complexity of the methodology used by Guan et al. renders comparisons with our methodology practically impossible.

Our work is not devoid of limitations and should be interpreted in the context of its retrospective and population-based design. First, the SEER database focuses on invasive UTUC, since T_{is} and T_a patients are not included. In consequence, our observations are based on more advanced stage and grade distribution and are not directly comparable with studies that used the entire UTUC population as reference. However, T_{is} and T_a patients should ideally not be treated with RNU. In consequence, their exclusion from SEER database does not represent an important limitation for studies that focus on RNU. Second, disease progression or disease recurrence data are not available in the SEER database. In consequence, they cannot be examined as endpoints. Third, the SEER database does not allow to ascertain either type or duration of chemotherapy. Fourth, due to the short median follow-up, future studies with longer follow-up should be

done to confirm or refuse our results. Fifth, our study did not benefit of central pathology review. Sixth, our analyses could not assess intra- and interobserver variability, which are essential in clinical practice. Finally, the SEER database represents a proportion of the United States populations. In consequence, our findings are only applicable to patients from the United States and are not generalizable to patients from other parts of the world. However, these limitations apply to this and to all other studies based on the SEER database.

Conclusion

From a statistical standpoint, either 1973 WHO or 2004/2016 WHO grading system improves the accuracy of CSM prediction to the same extent. In consequence, other considerations such as intra- and interobserver variability may represent additional metrics to consider in deciding which grading system is better.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10147-021-01941-9>.

Funding Open access funding provided by Università degli Studi di Napoli Federico II within the CRUI-CARE Agreement.

Declarations

Conflict of interest Our research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Siegel RL, Miller KD, Jemal A (2019) Cancer statistics, 2019. *CA Cancer J Clin* 69(1):7–34
2. Roupêt M, Babjuk M, Compérat E et al (2018) European association of urology guidelines on upper urinary tract urothelial carcinoma: 2017 update. *Eur Urol* 73(1):111–122
3. Collà Ruvolo C, Nocera L, Stolzenbach LF et al (2020) Incidence and survival rates of contemporary patients with invasive upper

- tract urothelial carcinoma. *Eur Urol Oncol*. <https://doi.org/10.1016/j.euo.2020.11.005>
4. Califano G, Xylinas E (2020) Re: Phase II trial of neoadjuvant systemic chemotherapy followed by extirpative surgery in patients with high grade upper tract urothelial carcinoma. *Eur Urol* 78(1):113–114
 5. Califano G, Ouzaid I, Verze P et al (2020) Immune checkpoint inhibition in upper tract urothelial carcinoma. *World J Urol*. <https://doi.org/10.1007/s00345-020-03502-7>
 6. Margulis V, Shariat SF, Matin SF et al (2009) Outcomes of radical nephroureterectomy: a series from the upper tract urothelial carcinoma collaboration. *Cancer* 115(6):1224–1233
 7. Lughezzani G, Burger M, Margulis V et al (2012) Prognostic factors in upper urinary tract urothelial carcinomas: a comprehensive review of the current literature. *Eur Urol* 62(1):100–114
 8. Mbeutcha A, Rouprêt M, Kamat AM et al (2017) Prognostic factors and predictive tools for upper tract urothelial carcinoma: a systematic review. *World J Urol* 35(3):337–353
 9. Petrelli F, Yasser Hussein MI, Vavassori I et al (2017) Prognostic factors of overall survival in upper urinary tract carcinoma: a systematic review and meta-analysis. *Urology* 100:9–15
 10. Fojecki G, Magnusson A, Traxer O et al (2019) Consultation on UTUC, Stockholm 2018 aspects of diagnosis of upper tract urothelial carcinoma. *World J Urol* 37(11):2271–2278
 11. Mostofi FK, Sobin LH, Torloni H et al (1973) Histological typing of urinary bladder tumours. World Health Organization. <https://apps.who.int/iris/handle/10665/41533>. Accessed Nov 4 2020
 12. Compérat EM, Burger M, Gontero P et al (2019) 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* 5(3):457–466
 13. Humphrey PA, Moch H, Cubilla AL et al (2016) The 2016 WHO classification of tumours of the urinary system and male genital organs—part B: prostate and bladder tumours. *Eur Urol* 70(1):106–119
 14. Miyamoto H, Miller JS, Fajardo DA et al (2010) Non-invasive papillary urothelial neoplasms: the 2004 WHO/ISUP classification system. *Pathol Int* 60(1):1–8
 15. Chen Z, Ding W, Xu K et al (2012) The 1973 WHO classification is more suitable than the 2004 WHO classification for predicting prognosis in non-muscle-invasive bladder cancer. *PLoS ONE* 7(10):e47199
 16. SEER Incidence Database—SEER Data & Software. SEER. <https://seer.cancer.gov/data/index.html>. Accessed 8 Oct 2020
 17. Rink M, Fajkovic H, Cha EK et al (2012) Death certificates are valid for the determination of cause of death in patients with upper and lower tract urothelial carcinoma. *Eur Urol* 61(4):854–855
 18. Collà Ruvolo C, Nocera L, Stolzenbach LF et al (2021) Tumor size predicts muscle-invasive and non-organ-confined disease in upper tract urothelial carcinoma at radical nephroureterectomy. *Eur Urol Focus*. <https://doi.org/10.1016/j.euf.2021.03.003>
 19. Park J, Ha SH, Min GE et al (2009) The protective role of renal parenchyma as a barrier to local tumor spread of upper tract transitional cell carcinoma and its impact on patient survival. *J Urol* 182(3):894–899
 20. Browne BM, Stensland KD, Moynihan MJ et al (2018) An analysis of staging and treatment trends for upper tract urothelial carcinoma in the national cancer database. *Clin Genitourin Cancer* 16(4):e743–e750
 21. Singla N, Fang D, Su X et al (2018) Preoperative predictors of nonorgan-confined disease in upper-tract urothelial carcinoma differ between China and the United States. *Urol Oncol Semin Orig Investig* 36(3):88.e11–88.e18
 22. Rouprêt M, Hupertan V, Traxer O et al (2006) Comparison of open nephroureterectomy and ureteroscopic and percutaneous management of upper urinary tract transitional cell carcinoma. *Urology* 67(6):1181–1187
 23. Collà Ruvolo C, Würnschimmel C, Wenzel M et al (2021) Comparison between 1973 and 2004/2016 WHO grading systems in patients with Ta urothelial carcinoma of urinary bladder. *J Clin Pathol*. <https://doi.org/10.1136/jclinpath-2021-207400>
 24. Soukup V, Čapoun O, Cohen D et al (2017) 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* 72(5):801–813
 25. van Rhijn BWG, van Leenders GJLH, Ooms BCM et al (2010) The pathologist’s mean grade is constant and individualizes the prognostic value of bladder cancer grading. *Eur Urol* 57(6):1052–1057
 26. May M, Brookman-Amisshah S, Roigas J et al (2010) Prognostic accuracy of individual uropathologists in noninvasive urinary bladder carcinoma: a multicentre study comparing the 1973 and 2004 World Health Organisation classifications. *Eur Urol* 57(5):850–858
 27. Mariappan P, Fineron P, O’Donnell M et al (2021) Combining two grading systems: the clinical validity and inter-observer variability of the 1973 and 2004 WHO bladder cancer classification systems assessed in a UK cohort with 15 years of prospective follow-up. *World J Urol* 39(2):425–431
 28. MacLennan GT, Kirkali Z, Cheng L (2007) Histologic grading of noninvasive papillary urothelial neoplasms. *Eur Urol* 51(4):889–898
 29. van de Putte EEF, Bosschieter J, van der Kwast TH et al (2018) The World Health Organization 1973 classification system for grade is an important prognosticator in T1 non-muscle-invasive bladder cancer. *BJU Int* 122(6):978–985
 30. Guan B, Tang S, Zhan Y et al (2019) Prognostic performance of the 1973 and 2004 WHO grading classification in upper tract urothelial carcinoma. *Urol Oncol Semin Orig Investig* 37(8):529.e19–529.e25

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.