Observational Study

Medicine®

OPEN

# Comparison between whole mount tissue preparations and virtual tissue microarray samples for measuring Ki-67 and apoptosis indices in human bladder cancer

## A cross-sectional study

Hisashi Oshiro, MD, PhD[a,d,f], Bogdan A. Czerniak, MD, PhD[a], Kentaro Sakamaki, PhD[b], Koji Tsuta, MD, PhD[a,e], Jolanta Bondaruk, PhD[a], Afsaneh Keyhani, MS[c], Colin P. Dinney, MD[c], Takeshi Nagai, MD[d], Ashish M. Kamat, MD[c,*]

**Abstract**

Recent tissue microarray (TMA)-based studies have shown that cell proliferation- and apoptosis-related biomarkers are associated with clinical outcomes in patients with bladder urothelial carcinoma. However, little is known about the differences in these biomarker measurements between whole mount tissue preparations and TMAs. This study aimed to elucidate the discrepancy in the measurements of Ki-67 indices (KIs) and apoptosis indices (AIs) between whole mount tissue preparations and TMAs of bladder urothelial carcinoma samples.

Whole mount tissue preparations for Ki-67 immunohistochemistry and terminal deoxynucleotidyl transferase dUTP nick end labeling were made from 30 patients who underwent transurethral resection of bladder urothelial carcinoma. Digital microscopy-assisted virtual TMAs, consisting of 3 small round areas (1 or 0.6 mm in diameter), were generated from the same whole mount tissue preparations. The measurement results in highly reactive areas of biomarkers were compared between the whole mount tissue preparation- and the TMA-based methods. Bland–Altman plot analysis, regression analysis, and Kendall $\tau$ were performed to investigate differences in the measurement results, systematic biases, and correlations between biomarkers.

Although the Bland–Altman plot analysis demonstrated that almost all the plots were within the limits of agreement, fixed biases were detected in the 1- and 0.6-mm TMAs for the KI (0.181 and 0.222, respectively) and the AI (0.055 and 0.063, respectively). Proportional biases were also detected in the 1- and 0.6-mm TMAs for the AI ($P < 0.001$ and $P < 0.001$, respectively). Furthermore, positive correlations between KIs and AIs were observed in whole mount tissue preparations ($r = 0.260$, $P = 0.044$) and in the 1 mm TMAs ($r = 0.375$, $P = 0.004$); however, no such correlation was observed in the 0.6 mm TMAs.

Our study suggests that the measurement results for certain biomarkers of bladder urothelial carcinoma obtained from TMA-based samples can be susceptible to systematic bias, and the lack of correlation between biomarkers cannot be avoided as it is in whole mount tissue preparations. Virtual TMAs can help identify systematic bias and establish a better sampling strategy prior to performing high-throughput TMAs for biomarker studies.

**Abbreviations:** AI = apoptosis index, KI = Ki-67 index, TMA = tissue microarray.

**Keywords:** Bland–Altman plot analysis, dimensional measurement accuracy, precision, systematic bias, terminal deoxynucleotidyl transferase dUTP nick end labeling, tissue microarray, urinary bladder cancer

*Editor: Farid Azmoudeh-Ardalan.*

*Funding/support: This research was supported by Grants-in-Aid for Specialized Programs of Research Excellence for bladder cancer research from the National Cancer Institute of the United States (to AMK and CPD) and by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan (no. 24590242 to HO and no. 24700501 to TN).*

*The authors have no conflicts of interest to disclose.*

*Supplemental Digital Content is available for this article.*

*[a] Department of Pathology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA, [b] Department of Biostatistics and Epidemiology, Yokohama City University Graduate School of Medicine, Yokohama, Kanagawa, Japan, [c] Department of Urology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA, [d] Department of Anatomic Pathology, Tokyo Medical University, Tokyo, [e] Department of Clinical Sciences and Laboratory Medicine, Kansai Medical University, Hirakata, Osaka, [f] Department of Pathology, Jichi Medical University, Shimotsuke, Tochigi, Japan.*

*[*] Correspondence: Ashish M. Kamat, Department of Urology, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030, USA (e-mail: akamat@mdanderson.org).*



*Medicine (2016) 95:31(e4500)*

*Received: 23 February 2016 / Received in final form: 13 June 2016 / Accepted: 13 July 2016*

*http://dx.doi.org/10.1097/MD.0000000000004500*

## 1. Introduction

Biomarkers may help stratify bladder cancer patients to provide them with appropriate therapeutic strategies. Recent studies have demonstrated that cell proliferation- and apoptosis-associated biomarkers are useful for predicting the clinical outcomes of patients with bladder urothelial carcinoma.[1] The cell proliferation marker Ki-67 is a nuclear antigen expressed in the S, G1, G2, and M phases of the cell cycle. A study of patients who underwent transurethral resection for nonmuscle invasive bladder carcinoma demonstrated that a high Ki-67 index (KI) was an independent risk factor for disease recurrence and progression using a KI cut-off of 25%.[2] Other studies showed that with a cut-off of 20%, a high KI was associated with advanced pathological stage, higher tumor grade, lymphovascular invasion, metastasis, disease recurrence, and stage-adjusted disease-specific mortality in patients with bladder urothelial carcinoma who underwent radical cystectomy.[3,4] A low apoptosis index (AI) was associated with worse local control or a lower survival rate in patients with bladder urothelial carcinoma.[5–7] Similarly, a previous study demonstrated that low expression of cleaved caspase-3, a key apoptosis marker, was associated with a low survival rate in patients with bladder carcinoma.[8]

However, some studies do not correspond with these results.[9–12] The use of different cut-off points in various studies has certain clinical implications, making study comparisons difficult.[13] The causes of such discrepancies are thought to be differences in sampling, fixation, antigen retrieval, cell counting, tissue section thickness, antibody concentration, and/or reaction time protocols.[1,14,15] Of these factors, careful attention must be paid to sampling methodology because most biomarker studies of bladder urothelial carcinoma are now based on tissue microarrays (TMAs).[16]

TMA studies for bladder urothelial carcinoma typically use 2 or 3 core samples (1 or 0.6 mm in diameter) per case.[16] If the TMA-based measurement results for biomarkers are sufficient for clinical interpretation, then we can replace whole mount tissue preparations with TMA samples or use the 2 methods interchangeably. One approach to explore this issue is a virtual TMA investigation, which is a simulation study that uses virtual cores constructed from images taken from whole mount tissue preparations and does not require the physical punching of actual holes in donor paraffin blocks.[17,18] Although previous biomarker studies using virtual TMAs have compared measurement results obtained using whole mount tissue preparations and TMAs for breast carcinoma[18] and peripheral T-cell lymphoma,[17] no adequate data exist to demonstrate that bladder urothelial carcinoma TMAs produce acceptable measurements compared with conventional whole mount tissue preparations.

The aim of the present study was to elucidate the degree of discrepancy in the KI and AI measurement results acquired from whole mount tissue preparations and virtual TMAs of human bladder urothelial carcinoma when each measurement was obtained using constant measurement criteria.

## 2. Materials and methods

### 2.1. Case selection

This retrospective study was approved by the Institutional Review Board of the University of Texas MD Anderson Cancer Center (LAB07-0420). The required sample size for this study was estimated as 30, which we thought adequate for our study purpose and statistical interpretation based on the literature.[19–22] A total of 30 patients (10 Ta, 10 T1, and 10 T2–T4 patients) were retrospectively and consecutively retrieved from the computerized database of MD Anderson Cancer Center between 1998 and 2008. For inclusion in the study, patients were required to have undergone transurethral resection of bladder urothelial carcinoma for the 1st time at MD Anderson Cancer Center without receiving systemic chemotherapy, immunotherapy, radiation therapy of the pelvic region, or any intravesical therapies, including Bacillus Calmette–Guerin vaccine, prior to initial transurethral resection. The exclusion criterion was unavailability of formalin-fixed, paraffin-embedded tissue samples from the initially resected bladder tumor. Board-certified urologists reviewed patients' medical charts and confirmed information regarding patient age, gender, clinical history, clinical imaging, and treatment for bladder cancer. Board-certified anatomic pathologists evaluated histological samples and confirmed patients' pathological diagnoses and tumor grades in accordance with standards in the literature.[23]

### 2.2. Histopathological slide preparation

Among the initially resected bladder urothelial carcinoma samples, 1 formalin-fixed, paraffin-embedded tissue block was selected per patient by board-certified anatomic pathologists. Each block was sectioned at a thickness of 4 μm, and sections were mounted onto silane-coated glass slides and used for immunohistochemistry (for KI) and terminal deoxynucleotidyl transferase dUTP nick end labeling (for AI).

Immunohistochemistry was performed on whole mount tissue sections using a Ki-67 antibody (SP6; Lab Vision; Fremont, CA; 1:200 dilution) and an automated immunostainer (Nemesis 8400; Biocare Medical; Concord, CA). Sections and appropriate positive and negative control samples were first subjected to antigen retrieval using citrate buffer (pH 6.0) for 20 minutes at 95 to 100 °C, followed by incubation with primary antibody for 60 minutes at room temperature and colorization using an avidin–biotin complex detection method and diaminobenzidine, with procedures conducted in accordance with the manufacturer's instructions and methods described in the literature.[24]

Terminal deoxynucleotidyl transferase dUTP nick end labeling was performed on whole mount tissue sections using an Apoptosis in situ Detection Kit (Wako; Osaka, Japan) with diaminobenzidine colorization; appropriate control samples were used, and procedures were performed in accordance with the manufacturer's instructions and methods described in the literature.[25,26]

### 2.3. Production of virtual TMAs

Each whole mount tissue preparation was used to compare different sampling methods for the KI and AI: a whole mount tissue preparation-based method; and a TMA-based method consisting of 3 small round areas (cores) of either 1 or 0.6 mm in diameter that were selected from the same whole mount tissue preparations, which we called "virtual TMAs."

First, we made a small round grid with Excel (Microsoft Japan; Tokyo, Japan), in which each small, round area was assigned a number. Second, we printed the grid onto a transparent film such that the diameter of each small, round area was 1 or 0.6 mm. Third, we placed the transparent film on each whole mount tissue preparation and divided the bladder cancer tissue into small, round areas. Finally, for each whole mount tissue preparation, we used a random number table to randomly select 3 small, round

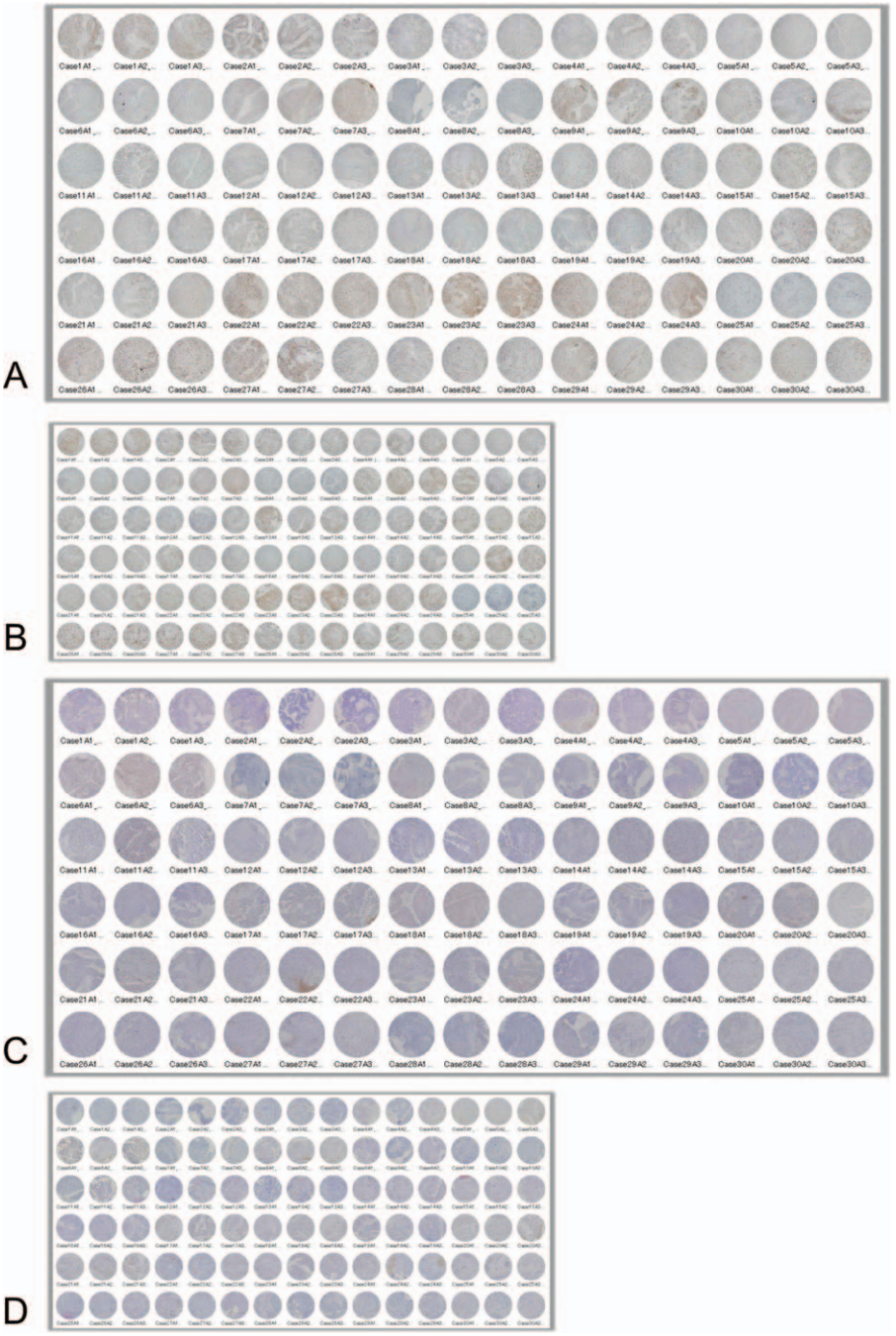


**Figure 1.** Virtual tissue microarrays used in this study. (A) Ninety core samples with a diameter of 1 mm (Ki-67 immunohistochemistry). (B) Ninety core samples with a diameter of 0.6 mm (Ki-67 immunohistochemistry). (C) Ninety core samples with a diameter of 1 mm (terminal deoxynucleotidyl transferase dUTP nick end labeling (TUNEL). (D) Ninety core samples with a diameter of 0.6 mm (TUNEL).

areas in which at least one third of the area contained neoplastic cells and then took digital photomicrographs. Thus, virtual TMAs were composed of samples with a diameter of 1 or 0.6 mm (Fig. 1 and Supplementary Figure 1, http://links.lww.com/MD/B183).

### *2.4. Histopathological image analysis*

All KI and AI measurements were performed twice by a board-certified anatomic pathologist (HO) with more than 15 years of experience. The rater measured the KI and AI of each image in the following order with an interval of at least 2 weeks between the different sample types: whole mount tissue preparations (as a reference standard), 1 mm TMA samples (as an index test), and 0.6 mm TMA samples (as an index test). The rater was blinded to the results of the reference standard and the index tests for the KI and AI until all measurements were completed.

We used widely accepted measurement protocols for KI and AI based on the literature.[3,5,6,27,28] Briefly, in the whole mount histological preparations and virtual TMA samples, highly reactive areas, which we called "hot spots," were observed under a light microscope using an imaging system (DP70; Olympus;

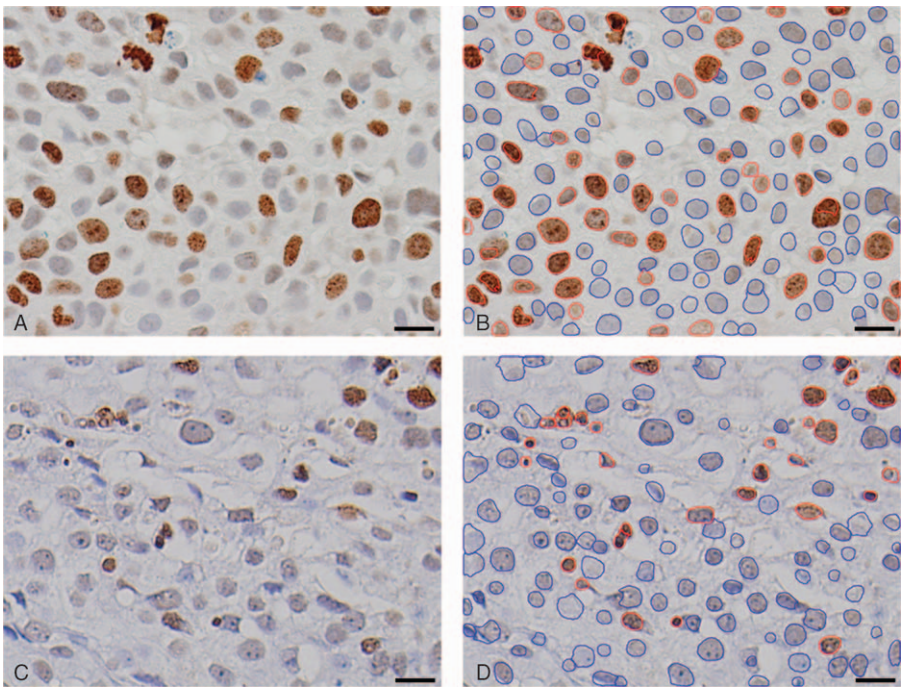


**Figure 2.** Photomicrographs of bladder cancer tissue used for digital microscope-assisted morphometry. (A) Ki-67 immunohistochemistry (40 × objective; scale bar: 20 μm). (B) Ki-67 immunohistochemistry analyzed by NuclearQuant (3DHISTECH; Budapest, Hungary). Ki-67-negative nuclei are outlined in blue, whereas Ki-67-positive nuclei are outlined in red (40 × objective; scale bar: 20 μm). (C) Terminal deoxynucleotidyl transferase dUTP nick end labeling (TUNEL) (40 × objective; scale bar: 20 μm). (D) TUNEL analyzed by NuclearQuant, TUNEL-negative nuclei are outlined in blue, whereas TUNEL-positive nuclei are outlined in red (40 × objective; scale bar: 20 μm).

Tokyo, Japan). A high-power field (215 μm × 160 μm = 3440 $\mu m^2$) image was digitally obtained, saved in tagged image file format (1020 × 768 pixels), and converted into MRXS files by Slide Converter (3DHISTECH; Budapest, Hungary). These digital images were viewed with the assistance of Pannoramic Viewer (3DHISTECH) and NuclearQuant (3DHISTECH) image analysis software under detailed measurement settings (Fig. 2). After reviewing, adjusting and confirming the computer graphic data, the most highly reactive images were selected to determine the KI and AI. The numbers of positive and total neoplastic cells counted in each image were added in descending order of the ratio of positive cells to total cells until the total number of neoplastic cells exceeded 999. If the 3 microarray samples did not contain more than 999 neoplastic cells, all the neoplastic cells in the 3 microarray samples were counted to determine the KI or AI. For WMTPs, we initially measured numerous potential hot-spot high-power field images (for the KI, an average of 12.5 images per case; for the AI, an average of 14.6 images per case). After measuring and ranking these images, we selected hot-spot images in descending order (for the KI, an average of 7.6 images per case; for the AI, an average of 7.0 images per case). For virtual TMAs, we initially measured many more potential hot-spot images (for the KI, an average of 14.5 images from 1 mm TMAs per case and 13.5 images from 0.6 mm TMAs per case; for the AI, an average of 14.0 images from 1 mm TMAs per case and 12.3 images from 0.6 mm TMAs per case) than the number of hot-spot high-power images that were eventually selected (for the KI, an average of 8.0 images from 1 mm TMAs per case and 7.6 images from 0.6 mm TMAs per case; for the AI, an average of 7.5 images from 1 mm TMAs per case and 7.6 images from 0.6 mm TMAs per case).

### 2.5. Statistical analyses

The 1st and 2nd measurement values were used to assess intrarater reliability, and the mean values were used for the Bland–Altman plot analysis and Kendall τ. For the assessment of intrarater reliability, the intraclass correlation coefficients for a single measurement and for the average of 2 measurements were investigated using IBM SPSS Statistics 21 (IBM Japan; Tokyo, Japan). Bland–Altman plot analysis[19,29,30] was used to compare the 2 measurement methods using MedCalc 14 (MedCalc; Ostend, Belgium). In addition, ordinary least-squares regression analysis was performed on the Bland–Altman plot using MedCalc 14. Fixed bias was indicated if the 95% confidence interval (CI) for the mean value of the difference did not contain 0. Proportional bias was indicated if the slope of the ordinary least regression of the differences in the means differed significantly from 0 ($P < 0.05$). Kendall τ was used to examine correlations between KIs and AIs in the whole mount tissue preparations and in the virtual TMA samples using IBM SPSS Statistics 21. In all statistical analyses, a $P$-value of less than 0.05 (2-sided) was considered to indicate statistical significance.

## 3. Results

A total of 30 patients (10 Ta, 10 T1, and 10 T2-T4 patients) fulfilled our inclusion criteria, and none of the patients were eliminated based on our exclusion criterion. The patients' characteristics are summarized in Table 1. Our study population had a mean age of 65.9 years (range, 45–89), the men to women ratio was 24:6, and 4 patients had minor histological components (glandular, micropapillary, or squamous differentiation) other than urothelial carcinoma.

**Table 1**

**Clinicopathological features of 30 cases with bladder carcinomas.**

| Case | Age | Sex | TNM (UICC) | Histology (grade) |
|---|---|---|---|---|
| 1 | 69 | Man | TaN0M0 | Urothelial carcinoma, high grade |
| 2 | 57 | Man | TaN0M0 | Urothelial carcinoma, high grade |
| 3 | 64 | Man | TaN0M0 | Urothelial carcinoma, high grade |
| 4 | 68 | Woman | TaN0M0 | Urothelial carcinoma, high grade |
| 5 | 83 | Man | TaN0M0 | Urothelial carcinoma, high grade |
| 6 | 57 | Man | TaN0M0 | Urothelial carcinoma, high grade |
| 7 | 64 | Man | TaN0M0 | Urothelial carcinoma, high grade |
| 8 | 73 | Man | TaN0M0 | Urothelial carcinoma, high grade |
| 9 | 81 | Man | TaN0M0 | Urothelial carcinoma, high grade |
| 10 | 62 | Man | TaN0M0 | Urothelial carcinoma, high grade |
| 11 | 71 | Man | T1N0M0 | Urothelial carcinoma, high grade, with focal glandular differentiation |
| 12 | 53 | Man | T1N0M0 | Urothelial carcinoma, high grade |
| 13 | 62 | Man | T1N0M0 | Urothelial carcinoma, high grade |
| 14 | 47 | Woman | T1N0M0 | Urothelial carcinoma, high grade |
| 15 | 81 | Woman | T1N0M0 | Urothelial carcinoma, high grade |
| 16 | 61 | Woman | T1N0M0 | Urothelial carcinoma, high grade |
| 17 | 72 | Man | T1N0M0 | Urothelial carcinoma, high grade |
| 18 | 72 | Man | T1N0M0 | Urothelial carcinoma, high grade |
| 19 | 50 | Man | T1N0M0 | Urothelial carcinoma, high grade |
| 20 | 65 | Man | T1N0M0 | Urothelial carcinoma, high grade |
| 21 | 52 | Man | T2N0M0 | Urothelial carcinoma, high grade |
| 22 | 88 | Man | T3N0M0 | Urothelial carcinoma, high grade |
| 23 | 62 | Man | T2N2M1 | Urothelial carcinoma, high grade, with focal micropapillary differentiation |
| 24 | 69 | Man | T2N1M0 | Urothelial carcinoma, high grade |
| 25 | 75 | Woman | T2N0M0 | Urothelial carcinoma, high grade |
| 26 | 52 | Man | T4N0M0 | Urothelial carcinoma, high grade, with focal squamous differentiation |
| 27 | 57 | Man | T3N2M0 | Urothelial carcinoma, high grade |
| 28 | 45 | Man | T2N0M0 | Urothelial carcinoma, high grade |
| 29 | 89 | Man | T2N0M0 | Urothelial carcinoma, high grade |
| 30 | 75 | Woman | T2N0M0 | Urothelial carcinoma, high grade, with focal squamous and glandular differentiation |

TMA = tissue microarray, UICC = Union for International Cancer Control.

The 1st and 2nd measurement results are shown in Table 2. The intraclass correlation coefficients for a single KI measurement and for the average of 2 KI measurements were 0.987 (95% CI, 0.972–0.994, $P < 0.001$) and 0.993 (95% CI, 0.986–0.997, $P < 0.001$) in the whole mount tissue preparations, 0.995 (95% CI, 0.989–0.997, $P < 0.001$) and 0.997 (95% CI, 0.994–0.999, $P < 0.001$) in the 1 mm TMA samples, and 0.991 (95% CI, 0.982–0.996, $P < 0.001$) and 0.996 (95% CI, 0.991–0.998, $P < 0.001$) in the 0.6 mm TMA samples, respectively. The intraclass correlation coefficients for a single AI measurement and for the average of 2 AI measurements were 0.989 (95% CI, 0.978–0.995, $P < 0.001$) and 0.995 (95% CI, 0.989–0.997, $P < 0.001$) in the whole mount tissue preparations, 0.992 (95% CI, 0.984–0.996, $P < 0.001$) and 0.996 (95% CI, 0.992–0.998, $P < 0.001$) in the 1 mm TMA samples, and 0.994 (95% CI, 0.989–0.997, $P < 0.001$) and 0.997 (95% CI, 0.994–0.999, $P < 0.001$) in the 0.6 mm TMA samples, respectively.

As shown in Fig. 3, the KIs measured from the whole mount tissue preparations were as follows: median, 0.502; minimum, 0.112; maximum, 0.821; and interquartile range, 0.302. The KIs measured using the 1 mm virtual TMA samples were as follows: median, 0.273; minimum, 0.026; maximum, 0.729; and interquartile range, 0.254. The KIs measured using the 0.6 mm virtual TMA samples were as follows: median, 0.265; minimum, 0.031; maximum, 0.739; and interquartile range, 0.247. Only case number 25 did not contain more than 999 neoplastic cells in the 0.6 mm TMA samples for the KI assessment (1st measurement, 94/711 = 0.132; 2nd measurement, 99/720 = 0.138). The other TMA samples all contained more than 1000 neoplastic cells.

The AIs measured in the whole mount tissue preparations were as follows: median, 0.102; minimum, 0.034; maximum, 0.248; and interquartile range, 0.010. The AIs measured using the 1 mm virtual TMA samples were as follows: median, 0.045; minimum, 0.011; maximum, 0.132; and interquartile range, 0.044. The AIs measured using the 0.6 mm virtual TMA samples were as follows: median, 0.044; minimum, 0.005; maximum, 0.122; and interquartile range, 0.031.

Figure 4A shows a comparison of the KIs obtained using whole mount tissue preparations and 1 mm TMA samples. Although all the plots were within the limits of agreement, a fixed bias was indicated by the mean difference of 0.181 (95% CI: 0.137–0.225). However, a proportional bias was not evident (slope: 0.065, $P = 0.594$).

Figure 4B shows a comparison of the KIs obtained using whole mount tissue preparations and 0.6 mm TMA samples. Although nearly all the plots were within the limits of agreement, a fixed bias was indicated by the mean difference of 0.222 (95% CI: 0.176–0.268), but a proportional bias was not evident (slope: 0.256, $P = 0.555$).

Figure 4C shows a comparison of the AIs obtained using whole mount tissue preparations and virtual 1 mm TMA samples. In this comparison, although almost all the plots were within the limits of agreement, a fixed bias was indicated by the mean difference of

**Table 2**

**Ki-67 and apoptosis indices in bladder carcinomas.**

| | Ki-67 index | | | | | | | | | Apoptosis index | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Whole mount tissue preparation | | | 1.0-mm diameter, virtual tissue microarray (3 cores) | | | 0.6-mm diameter, virtual tissue microarray (3 cores) | | | Whole mount tissue preparation | | | 1-mm diameter, virtual tissue microarray (3 cores) | | | 0.6-mm diameter, virtual tissue microarray (3 cores) | | |
| Case | 1st measurement | 2nd measurement | Mean | 1st measurement | 2nd measurement | Mean | 1st measurement | 2nd measurement | Mean | 1st measurement | 2nd measurement | Mean | 1st measurement | 2nd measurement | Mean | 1st measurement | 2nd measurement | Mean |
| 1 | 62.16% | 63.14% | 62.65% | 46.98% | 44.40% | 45.69% | 52.08% | 50.57% | 51.32% | 14.18% | 14.88% | 14.53% | 4.11% | 3.67% | 3.89% | 1.49% | 1.49% | 1.49% |
| 2 | 52.53% | 52.81% | 52.67% | 43.83% | 44.43% | 44.13% | 30.23% | 32.00% | 31.12% | 5.01% | 4.94% | 4.97% | 3.85% | 4.26% | 4.06% | 5.04% | 4.78% | 4.91% |
| 3 | 26.57% | 27.88% | 27.23% | 21.73% | 22.02% | 21.87% | 19.34% | 19.57% | 19.46% | 5.16% | 4.85% | 5.01% | 3.17% | 3.85% | 3.51% | 2.24% | 2.41% | 2.32% |
| 4 | 51.44% | 53.90% | 52.67% | 32.73% | 33.58% | 33.16% | 21.25% | 23.60% | 22.42% | 19.14% | 19.33% | 19.23% | 7.09% | 7.10% | 7.09% | 4.72% | 4.62% | 4.67% |
| 5 | 45.60% | 47.69% | 46.65% | 13.73% | 13.54% | 13.63% | 11.44% | 13.36% | 12.40% | 24.69% | 24.89% | 24.79% | 2.24% | 2.49% | 2.36% | 10.16% | 9.66% | 9.91% |
| 6 | 21.96% | 22.26% | 22.11% | 5.89% | 7.15% | 6.52% | 6.87% | 8.13% | 7.50% | 4.12% | 3.83% | 3.97% | 2.24% | 2.32% | 2.28% | 3.63% | 3.07% | 3.35% |
| 7 | 43.61% | 44.67% | 44.14% | 14.88% | 16.37% | 15.62% | 11.92% | 12.21% | 12.07% | 15.43% | 15.59% | 15.51% | 3.73% | 3.27% | 3.50% | 3.03% | 3.17% | 3.10% |
| 8 | 10.91% | 11.45% | 11.18% | 2.03% | 3.07% | 2.55% | 4.66% | 5.44% | 5.05% | 8.65% | 9.23% | 8.94% | 4.51% | 4.65% | 4.58% | 4.93% | 4.95% | 4.94% |
| 9 | 48.27% | 45.92% | 47.10% | 45.00% | 43.20% | 44.10% | 37.93% | 34.21% | 36.07% | 3.69% | 3.85% | 3.77% | 2.38% | 2.80% | 2.59% | 2.95% | 3.05% | 3.00% |
| 10 | 83.71% | 80.46% | 82.08% | 59.29% | 57.85% | 58.57% | 32.53% | 33.56% | 33.05% | 21.25% | 22.53% | 21.89% | 8.90% | 9.05% | 8.98% | 10.41% | 10.56% | 10.48% |
| 11 | 32.71% | 34.30% | 33.50% | 22.34% | 22.28% | 22.31% | 19.20% | 23.73% | 21.47% | 6.52% | 7.56% | 7.04% | 4.84% | 5.60% | 5.22% | 6.19% | 6.77% | 6.48% |
| 12 | 38.19% | 36.80% | 37.49% | 10.55% | 11.32% | 10.93% | 13.07% | 13.94% | 13.51% | 4.99% | 5.31% | 5.15% | 1.08% | 1.14% | 1.11% | 0.46% | 0.46% | 0.46% |
| 13 | 73.30% | 72.88% | 73.09% | 41.28% | 40.36% | 40.82% | 42.20% | 42.19% | 42.20% | 14.90% | 15.18% | 15.04% | 12.91% | 13.56% | 13.23% | 4.23% | 4.64% | 4.43% |
| 14 | 33.65% | 31.28% | 32.46% | 22.90% | 21.22% | 22.06% | 14.65% | 17.92% | 16.28% | 9.92% | 10.20% | 10.06% | 5.88% | 5.44% | 5.66% | 5.54% | 5.51% | 5.53% |
| 15 | 38.71% | 33.55% | 36.13% | 25.27% | 24.39% | 24.83% | 28.87% | 29.01% | 28.94% | 12.12% | 12.45% | 12.28% | 5.87% | 6.09% | 5.98% | 4.80% | 5.24% | 5.02% |
| 16 | 62.72% | 60.03% | 61.38% | 21.53% | 20.71% | 21.12% | 17.18% | 18.37% | 17.78% | 3.23% | 3.49% | 3.36% | 3.04% | 3.48% | 3.26% | 1.45% | 1.46% | 1.46% |
| 17 | 63.68% | 60.78% | 62.23% | 23.57% | 23.93% | 23.75% | 21.82% | 23.15% | 22.49% | 4.95% | 5.31% | 5.13% | 4.32% | 4.64% | 4.48% | 2.61% | 2.92% | 2.77% |
| 18 | 16.96% | 17.31% | 17.13% | 9.87% | 9.84% | 9.86% | 3.16% | 3.11% | 3.13% | 5.06% | 5.55% | 5.31% | 3.04% | 2.78% | 2.91% | 1.46% | 1.83% | 1.64% |
| 19 | 47.60% | 47.97% | 47.78% | 21.62% | 25.00% | 23.31% | 19.96% | 20.41% | 20.18% | 6.16% | 7.03% | 6.59% | 4.51% | 4.69% | 4.60% | 3.34% | 3.49% | 3.42% |
| 20 | 74.77% | 68.84% | 71.80% | 39.13% | 40.31% | 39.72% | 40.65% | 40.62% | 40.63% | 25.05% | 24.54% | 24.80% | 12.73% | 13.35% | 13.04% | 11.96% | 12.14% | 12.05% |
| 21 | 61.00% | 58.31% | 59.66% | 32.30% | 27.38% | 29.84% | 31.43% | 35.61% | 33.52% | 15.77% | 14.96% | 15.36% | 12.20% | 11.24% | 11.72% | 6.65% | 6.21% | 6.43% |
| 22 | 64.43% | 62.83% | 63.63% | 59.14% | 56.12% | 57.63% | 40.02% | 40.77% | 40.39% | 11.50% | 11.36% | 11.43% | 8.72% | 8.23% | 8.47% | 5.69% | 5.99% | 5.84% |
| 23 | 79.89% | 77.55% | 78.72% | 74.33% | 71.51% | 72.92% | 74.01% | 73.81% | 73.91% | 13.30% | 15.99% | 14.64% | 7.78% | 7.88% | 7.83% | 12.04% | 12.45% | 12.24% |
| 24 | 82.34% | 78.94% | 80.64% | 57.31% | 54.57% | 55.94% | 41.13% | 39.50% | 40.31% | 4.39% | 4.17% | 4.28% | 4.22% | 4.23% | 4.22% | 4.18% | 3.70% | 3.94% |
| 25 | 21.29% | 17.06% | 19.18% | 17.36% | 15.37% | 16.36% | 13.22% | 13.75% | 13.49% | 6.92% | 5.88% | 6.40% | 3.49% | 3.45% | 3.47% | 3.00% | 3.48% | 3.24% |
| 26 | 54.20% | 51.62% | 52.91% | 44.21% | 43.06% | 43.64% | 40.60% | 39.80% | 40.20% | 12.49% | 12.26% | 12.38% | 2.27% | 2.54% | 2.41% | 2.14% | 2.08% | 2.11% |
| 27 | 77.07% | 69.99% | 73.53% | 72.69% | 72.96% | 72.83% | 49.85% | 48.11% | 48.98% | 18.60% | 20.47% | 19.54% | 10.82% | 11.10% | 10.96% | 8.33% | 8.65% | 8.49% |
| 28 | 75.53% | 69.34% | 72.43% | 40.70% | 41.18% | 40.94% | 43.20% | 41.24% | 42.22% | 8.76% | 9.57% | 9.16% | 3.30% | 3.35% | 3.33% | 2.95% | 3.42% | 3.18% |
| 29 | 47.98% | 41.57% | 44.78% | 18.83% | 20.20% | 19.51% | 24.88% | 23.40% | 24.14% | 11.41% | 13.25% | 12.33% | 7.81% | 7.61% | 7.71% | 4.37% | 4.50% | 4.44% |
| 30 | 47.36% | 42.66% | 45.01% | 37.85% | 33.10% | 35.47% | 29.02% | 34.49% | 31.76% | 9.43% | 11.40% | 10.42% | 5.82% | 6.24% | 6.03% | 4.29% | 4.67% | 4.48% |

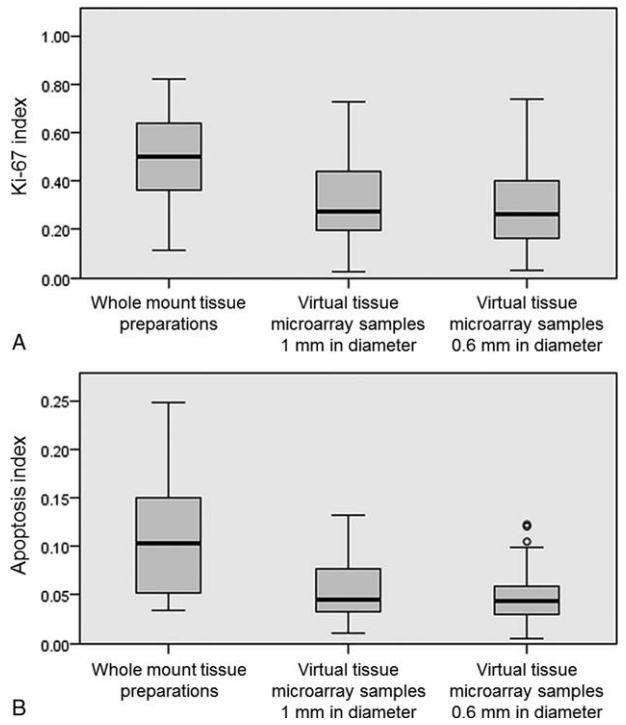


**Figure 3.** Boxplots of KI and AI. (A) KI measured using WMTPs and virtual TMA samples with diameters of 1 and 0.6 mm. (B) AI measured using WMTPs and virtual TMA samples with diameters of 1 and 0.6 mm. AI = apoptosis index, KI = Ki-67 index, TMA = tissue microarray, WMTP = whole mount tissue preparation.

0.055 (95% CI: 0.036–0.074), and a proportional bias was detected (slope: 0.772, $P < 0.001$).

Figure 4D shows a comparison of the AIs obtained using whole mount tissue preparations and the 0.6 mm TMA samples. Although all the plots were within the limits of agreement, a fixed bias was indicated by the mean difference of 0.063 (95% CI: 0.045–0.080), and a proportional bias was detected (slope: 0.798, $P < 0.001$).

Scatter plots to investigate the correlation between KIs and AIs are shown in Fig. 5. A positive correlation between KIs and AIs was observed in the whole mount tissue preparations ($r = 0.260$, $P = 0.044$) and the 1 mm virtual TMA samples ($r = 0.375$, $P = 0.004$); however, a correlation was not observed in the 0.6 mm virtual TMA samples ($r = 0.200$, $P = 0.121$).

## 4. Discussion

In the present study, the intraclass correlation coefficients for a single measurement and for an average of 2 measurements for the KI and AI were sufficient to support intrarater reliability according to the proposed criteria.[31] In the Bland–Altman plot analysis, differences between the KI and AI measurements in the TMA samples and whole mount tissue preparations were nearly all within the limits of agreement or the 95% CIs of these limits. However, we found fixed biases in the KI measurements and fixed and proportional biases in the AI measurements. A smaller TMA sample diameter was correlated with a larger systematic bias.[32] Our findings clearly demonstrated that these indices differed when they were measured using virtual TMA samples or whole mount tissue preparations.

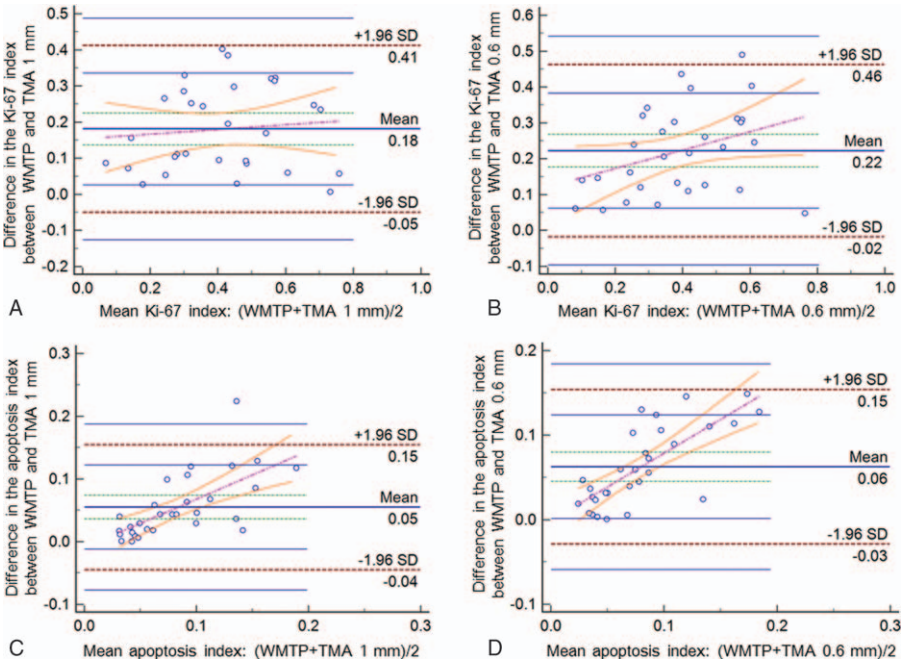


**Figure 4.** Bland–Altman plots for the comparison of different measurement methods. Each plot shows the differences between the 2 methods against the averages of the 2 methods; the lines represent the mean differences, 95% CI of the mean differences, upper and lower limits of agreement (mean differences ± 1.96 SD), 95% CI of the upper and lower limits of agreement, and regression of the differences. (A) Comparison of the KI measured using WMTPs or virtual TMA samples with a diameter of 1 mm (TMA 1 mm). (B) Comparison of the KI measured using WMTPs or virtual TMA samples with a diameter of 0.6 mm (TMA 0.6 mm). (C) Comparison of the AI measured using WMTPs or virtual TMA samples with a diameter of 1 mm (TMA 1 mm). (D) Comparison of the AI measured using WMTPs or virtual TMA samples with a diameter of 0.6 mm (TMA 0.6 mm). AI = apoptosis index, CI = confidence interval, KI = Ki-67 index, SD = standard deviation, TMA = tissue microarray, WMTP = whole mount tissue preparation.

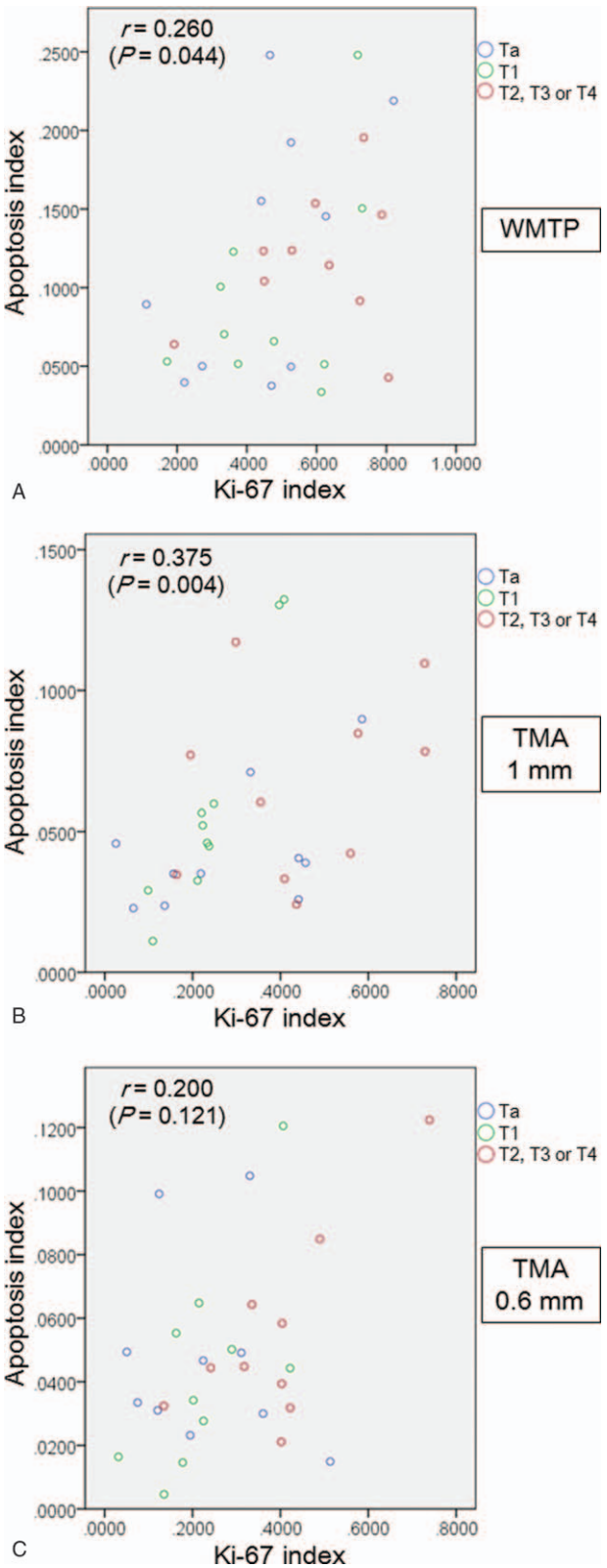


**Figure 5.** Correlation between KI and AI in WMTPs (A), virtual TMA samples with a diameter of 1 mm (B), and virtual TMA samples with a diameter of 0.6 mm (C). AI = apoptosis index, KI = Ki-67 index, TMA = tissue microarray, WMTP = whole mount tissue preparation.

There are 2 main types of error that interfere with research inferences: random error and systematic error. Random error is a wrong result due to chance; sources of variation are equally likely to distort estimates in the study in any direction.[32] Among several techniques for reducing the influence of random error, the simplest is to increase the sample size.[32] Systematic error is a wrong result due to bias; sources of variation distort the study findings in 1 direction.[32] Increasing the sample size has no effect on reducing systematic error.[32] Systematic error between different sampling methods can represent a fixed and/or a proportional bias. The only way to improve the accuracy of the estimate is to design the study in such a way that either reduces the size of the various biases or provides some information about them.[32] Although ordinary least-squares regression analysis is used in Bland–Altman plots to estimate the slope, caution is required when evaluating the fixed bias if a proportional bias exists. If a proportional bias exists ($slope \neq 0$) and the mean difference almost inevitably deviates from zero, there is a risk that fixed bias will be overestimated.[30] If a proportional bias exists in 1 direction (e.g., $slope > 0$) and a fixed bias exists in the opposite direction, there is a risk that the fixed bias will be underestimated.[33]

In the present study, the mean difference in the KI deviated considerably from zero (18.1% in the 1 mm TMA samples and 22.2% in the 0.6 mm TMA samples). Regarding the AI measurements, we found both fixed and proportional biases, although the fixed bias might have been overestimated due to the presence of a proportional bias. Hence, special attention should be paid to assigning cut-off points for clinical applications when using data obtained from different samplings of TMAs.[1–12,14,34]

We observed a positive correlation between KIs and AIs in whole mount tissue preparations and 1 mm virtual TMA samples; to the best of our knowledge, this had not yet been reported for human bladder cancer tissue. However, there was no correlation in the 0.6 mm microarray samples, indicating that smaller samples can result in a failure to detect a potentially important relationship between certain biomarkers. The likely cause of this phenomenon is that smaller samples are influenced by heterogeneous biomarker expression; therefore, much greater numbers of cores from sites containing adequate tumor cells would likely be required to obtain concordant results for WMTPs, 0.6 mm TMAs, and 1 mm TMAs. Because heterogeneity in expression appears to depend on the characteristics of the biomarker itself and the cells that are tested, such a discordance will be minimized when using homogeneously expressed biomarkers in TMA studies. If used carefully in this regard, TMA studies still have merit in that they are a low-cost, high-throughput method for determining the necessity for future detailed investigations between biomarkers and clinical outcomes.[16]

There are several limitations to the present study. First, we utilized only 1 counting protocol, as described in the methods section. Factors influencing measurement results by a single rater in this type of biomarker study may include the following: observational field areas depending on the magnification (e.g., 200 × or 400 ×); total number of neoplastic cells counted (e.g., 500 or 1000 cells); regions of interest (e.g., hot spots or random spots); (4) instruments used for observation (e.g., visual count under the microscope only or visual count in combination with a digital microscope-assisted image analyzer); and (5) the rater's measurement experience. Second, only 1 rater measured the KIs and AIs; therefore, interrater variability was not elucidated. Nevertheless, the present study provides important information about sampling biases inherent to TMAs. Because measurement results can significantly differ when smaller samples are included in TMAs, interrater variability will be an issue even when a

constant measurement method is employed by highly experienced pathologists. Further explorations will be necessary to identify a better sampling and counting strategy for biomarker studies that can minimize inter- and intrarater variability and precisely predict clinical outcomes.

In conclusion, our study suggests that KI and AI measurement results for TMA-based bladder urothelial carcinoma samples can be susceptible to systematic bias, and the lack of correlation between biomarkers cannot be avoided as it is with whole mount tissue preparations. Virtual TMAs in combination with Bland–Altman plot analysis can be useful for identifying a systematic bias before actually puncturing tissue blocks, and this will help establish a better sampling strategy for high-throughput TMA-based biomarker studies.

## Acknowledgments

The authors thank Junzo Kawaguchi, President of E-Path Co., Kanagawa, Japan, for providing useful advice on digital microscope-assisted morphometry.

## References

[1] Cheng L, Davison DD, Adams J, et al. Biomarkers in bladder cancer. Translational and clinical implications. Crit Rev Oncol Hematol 2014;89:73–111.
[2] Ding W, Gou Y, Sun C, et al. Ki-67 is an independent indicator in non-muscle invasive bladder cancer (NMIBC); combination of EORTC risk scores and Ki-67 expression could improve the risk stratification of NMIBC. Urol Oncol 2014;32:42.e13–9.
[3] Margulis V, Shariat SF, Ashfaq R, et al. Ki-67 is an independent predictor of bladder cancer outcome in patients treated with radical cystectomy for organ-confined disease. Clin Cancer Res 2006;12:7369–73.
[4] Margulis V, Lotan Y, Karakiewicz PI, et al. Multi-institutional validation of the predictive value of Ki-67 labeling index in patients with urinary bladder cancer. J Natl Cancer Inst 2009;101:114–9.
[5] Harada S, Sato R, Nakamura R, et al. The correlation between spontaneous and radiation-induced apoptosis in T3B bladder cancer (histological grade G3), and the precedence between the two kinds of apoptosis for predicting clinical prognosis. Int J Radiat Oncol Biol Phys 2000;48:1059–67.
[6] Rödel C, Grabenbauer GG, Rödel F, et al. Apoptosis, p53, bcl-2, and Ki-67 in invasive bladder carcinoma: possible predictors for response to radiochemotherapy and successful bladder preservation. Int J Radiat Oncol Biol Phys 2000;46:1213–21.
[7] Moonen L, Ong F, Gallee M, et al. Apoptosis, proliferating and p53, cyclin D1, and retinoblastoma gene expression in relation to radiation response in transitional cell carcinoma of the bladder. Int J Radiation Oncology Biol Phys 2001;49:1305–10.
[8] Karam JA, Lotan Y, Karakiewicz PI, et al. Use of combined apoptosis biomarkers for prediction of bladder cancer recurrence and mortality after radical cystectomy. Lancet Oncol 2007;8:128–36.
[9] Lara PC, Perez S, Rey A, et al. Apoptosis in carcinoma of the bladder: relation with radiation treatment results. Int J Radiat Oncol Biol Phys 1999;43:1015–9.
[10] Korkolopoulou P, Konstantinidou AE, Christodoulou P, et al. Apoptosis in bladder carcinomas detected with monoclonal antibody to single-stranded DNA: relation to cell cycle regulators and survival. Urology 2000;56:516–20.
[11] Behnsawy HM, Miyake H, Abdalla MA, et al. Expression of cell cycle-associated proteins in non-muscle-invasive bladder cancer: correlation with intravesical recurrence following transurethral resection. Urol Oncol 2011;29:495–501.
[12] Park J, Song C, Shin E, et al. Do molecular biomarkers have prognostic value in primary T1G3 bladder cancer treated with bacillus Calmette-Guerin intravesical therapy? Urol Oncol 2013;31:849–56.
[13] Grossman HB. Are biomarkers for bladder cancer beneficial? J Urol 2010;183:11–2.
[14] Rosenblatt R, Jonmarker S, Lewensohn R, et al. Current status of prognostic immunohistochemical markers for urothelial bladder cancer. Tumour Biol 2008;29:311–22.
[15] Sato M, Kojima M, Nagatsuma AK, et al. Optimal fixation for total preanalytic phase evaluation in pathology laboratories. A comprehensive study including immunohistochemistry, DNA, and mRNA assays. Pathol Int 2014;64:209–16.
[16] Kramer MW, Merseburger AS, Hennenlotter J, et al. Tissue microarrays in clinical urology – technical considerations. Scan J Urol Nephrol 2007;41:478–84.
[17] Pedersen MB, Riber-Hansen R, Nielsen PS, et al. Digital pathology for the validation of tissue microarrays in peripheral T-cell lymphomas. Appl Immunohistochem Mol Morphol 2014.
[18] Quintayo MA, Starczynski J, Yan FJ, et al. Virtual tissue microarrays: a novel and viable approach to optimizing tissue microarrays for biomarker research applied to ductal carcinoma in situ. Histopathology 2014;65:2–8.
[19] Altman DG, Bland JM. Measurement in medicine: the analysis of method comparison studies. Statistician 1983;32:307–18.
[20] Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307–10.
[21] Alvarez M, Pina DR, de Oliveira M, et al. Objective CT-based quantification of lung sequelae in treated patients with paracoccidioidomycosis. Medicine (Baltimore) 2014;93:e167.
[22] Giacomini G, Miranda JR, Pavan AL, et al. Quantification of pulmonary inflammatory processes using chest radiography: tuberculosis as the motivating application. Medicine (Baltimore) 2015;94:e1044.
[23] Montironi R, Lopez-Beltran A. The 2004 WHO classification of bladder tumors: a summary and commentary. Int J Surg Pathol 2005;13:143–53.
[24] Ekholm M, Beglerbegovic S, Grabau D, et al. Immunohistochemical assessment of Ki67 with antibodies SP6 and MIB1 in primary breast cancer: a comparison of prognostic value and reproducibility. Histopathology 2014;65:252–60.
[25] Gavrieli Y, Sherman Y, Ben-Sasson SA. Identification of programmed cell death in situ via specific labeling of nuclear DNA fragmentation. J Cell Biol 1992;119:493–501.
[26] Funaki N, Sasano H, Shizawa S, et al. Apoptosis and cell proliferation in biliary atresia. J Pathol 1998;186:429–33.
[27] Quintero A, Alvarez-Kindelan J, Luque RJ, et al. Ki-67 MIB1 labelling index and the prognosis of primary TaT1 urothelial cell carcinoma of the bladder. J Clin Pathol 2006;59:83–8.
[28] Gonzalez-Campora R, Davalos-Casanova G, Beato-Moreno A, et al. Apoptotic and proliferation indexes in primary superficial bladder tumors. Cancer Lett 2006;242:266–72.
[29] Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res 1999;8:135–60.
[30] Ludbrook J. Statistical technique for comparing measures and methods of measurement. A critical review. Clin Exp Pharmacol Physiol 2002;29:527–36.
[31] Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.
[32] Hulley SB, Newman TB, Cummings SR. Hulley SB, Cummings SR, Browner WS, et al. Getting started: the anatomy and physiology of clinical research. Designing Clinical Research. An Epidemiologic Approach 3rd ed.Philadelphia, PA:Lippincott Williams & Wilkins; 2007. 3–16.
[33] Ludbrook J. Confidence in Altman-Bland plots: a critical review of the method of differences. Clin Exp Pharmacol Physiol 2010;37:143–9.
[34] Luo Y, Zhang X, Mo M, et al. High Ki-67 immunohistochemical reactivity correlates with poor prognosis in bladder carcinoma: a comprehensive meta-analysis with 13,053 patients involved. Medicine (Baltimore) 2016;95:e3337.