

Doctoral (PhD) thesis

**Automated image cytometry for diagnosing  
bladder cancer by detecting genetically aberrant  
cells from voided urine via molecular cytogenetics**

Gábor Pajor

*PhD School: Interdisciplinary Medical Sciences*

*Head of PhD School: Professor Balázs Sümegei MD, PhD*

*PhD Program: Human Medical Genetics*

*Head of PhD Program; Supervisor: Professor Béla Melegh MD, PhD*

University of Pécs Medical School,

Department of Medical Genetics

Pécs, 2012

## List of Abbreviations

---

CEP	chromosome enumeration probe
CCD	charge-coupled device
CK	cytokeratin
CTC	circulating tumor cell
Del	deletion
DAPI	4,6-diamidine, 2-phenylindole dihydrochloride
ELISA	enzyme-linked immunosorbent assay
FISH	fluorescence in situ hybridization
FP	false positivity
FN	false negativity
HMWCK	high molecular weight cytokeratin
ICC	immuncytochemistry
LMWCK	low molecular weight cytokeratin
Mic	mean intensity within contour
MRD	minimal residual disease
PULMP	papillary urothelial neoplasm of low malignant potential
SLM	scanning light microscopy
SD	standard deviation
Tic	total intensity within contour
YAC	yeast arteficial chromosome

## Introduction

---

### Epidemiology and classification of bladder cancer

---

Bladder cancer is the fourth and seventh most common cancer amongst men and women respectively. Its occurrence in women is much lower than in men (1/3, 1/4), and it is even lower in women having gone through pregnancy. In 2008 there were 385,000 new patients diagnosed worldwide, while in the same year more than 150,000 patients lost their life because of bladder cancer. The territorial occurrence of the disease is very heterogeneous and there could be a variation in incidence as high as 14-fold, internationally; the highest incidence rates are found in Europe, North America and North Africa. The average death rate is 2-10 deaths/100,000/year; the average age in time of diagnosis is 65-70.

The most common known risk factors of bladder tumors in the Western world are cigarette smoking, and occupational exposure to urothelial carcinogens. The prior, being the most important risk element, accounts for 50% of cases in men and 35% in women. Cigarette smokers have a 2- to 4-fold increased risk of bladder cancer compared to non-smokers; this risk rises with increasing intensity and/or duration of smoking. Occupational exposure to urothelial carcinogens is the second most important risk factor, accounting for 5-20% of all bladder cancers; amongst others, exposure to aromatic amines (e.g. o-toluidin, benzidine) often used in rubber and dye industries have been associated with the development of bladder cancer. Other environmental factors include chronic urinary tract infections and exposure to radiation therapy. According to recent findings there is an increased incidence of bladder cancer in men with prostate cancer treated with radiotherapy. Opposed to Western world countries, chronic infection with schistosoma hematobium in developing countries, particularly in Africa and the Middle East, accounts for about 50% of the total burden. Thus far, there is no convincing evidence of inherited forms; nevertheless, a recent thorough segregation analysis could not exclude the possibility of an inherited subtype of bladder cancer.

Bladder cancer is a heterogenic entity and its final outcome moves on a wide scale. On one end of the spectrum lie the well differentiated early stage tumors, showing low progression rates; with initial endoscopic treatment and thorough follow up, these tumors are rarely life threatening. On the other end of the spectrum, we can find the poorly differentiated high grade high stage tumors, which are difficult to treat; these tumors are featured with a high rate of cancer mortality. Ninety percent of all bladder cancers are transitional cell carcinoma (TCC), 5% are squamous cell carcinoma while adenocarcinoma accounts for 1%. The histopathological grading of bladder cancer, according to the 1973 classification of the World Health Organization (WHO), is Grade 1-2-3. This system was changed by the WHO

and the International Society of Urological Pathology (ISUP) in 1998, which guideline was changed slightly again in 2004. The new system, however, had been criticized several times; moreover, in an official guideline the International Bladder Cancer Group (IBCG) suggested that both the 1973 and 1998 classification should remain in use until the latter does not receive sufficient confirmation via a large validation study. Taking all these circumstances into account, throughout this study we have used the original, and still prevalent G1-2-3 system. Staging was performed according to the Tumor-Node-Metastasis (TNM) system.

Of all TCCs 70% are superficial non-muscle invasive cancer (Ta, T1 and Tis stages); 50-70% of these tumors recur, of which 10-20% will progress to muscularis propria invasive disease (T2-4). To predict which patients will progress from superficial to invasive disease remains a challenge. Patients with low grade Ta tumors show 95% possibility of 15 year progression free survival, with practically no cancer mortality. Patients with high grade Ta tumors have a progression-free survival of 61% and a disease-specific survival of 74%, whereas patients with T1 disease have a progression free survival of 44% and a disease-specific survival of 62%.

The most common symptom of bladder cancer occurring in 85% of the cases is macroscopic but usually painless hematuria. However, it is important to note, that in only 10% of the cases will further examinations confirm the presence of a bladder carcinoma. Further, less common symptoms are frequent and painful voiding, renal pain, fever and anuria.

## Diagnosing bladder cancer

In case of suspicion for bladder cancer (e.g. hematuria), voided urine cytology, cystoscopy and radiological examination of the upper urinary tract remain the gold standard diagnostic procedures. Both cystoscopy and cytology have numerous advantages; however both are featured with some frequently noted drawbacks. Cystoscopy has high specificity and in most cases it is also featured with low false positivity, i.e high specificity values. However, some *in situ* lesions are known to remain undetected, moreover, the procedure is invasive and could cause infections and trauma. Another feature of cystoscopy is that it requires considerable utilization of resources and is often cited, along with the extremely high recurrence rate, as a driver of the very high cost of bladder cancer treatment; in fact, according to several calculations, the cost per patient with bladder cancer from diagnosis to death is the highest amongst all cancers.

One of the biggest advantage of voided urine cytology is that it does not require patient preparation or expensive equipment. It also has a high specificity (values even reach 100% in many studies), therefore the method is ideal for determining the absence of a bladder tumor, and false-positive interpretations are unusual. On the other hand it is much less efficient in detecting the presence of low-grade tumors, where the false-negative rate is higher. A wide

range of numbers appears in the literature, but most papers report sensitivities of 35% to 65%, sometimes even as low as 20%. The difficulty in identification of the disaggregated cells of low-grade urothelial neoplasms is almost certainly because of the absence of anaplastic changes in these tumors. Another frequently mentioned drawback of the method that its accuracy is very much dependent on the pathologist's expertise.

Because of the above mentioned drawbacks of cystoscopy and cytology there is a great need for a non-invasive, highly accurate, easily reproducible diagnostic assay. Conclusively, in the recent years numerous diagnostic tools have been developed, which can be classified into two main groups: cell based and soluble markers. Examples for the latter are (i) hematuria tests (ii) BTA stat and BTA-TRAK tests, (iii) nuclear matrix protein (NMP-22) assay, (iv) mitotic protein tests, (v) antiapoptotic protein assays, (vi) investigations of transcription factors, hialuronic acid. Cell based markers are, for example, (i) microsatellite analysis, (ii) telomerase tests, (iii) quantitative karyometry, and molecular cytogenetics (FISH).

### Cytogenetics of bladder cancer

The karyotypes reported in conjunction with bladder cancer in literature range from simple to very complex, and to date, no specific chromosomal aberrations had been described. One frequent abnormality is the monosomy of chromosome 9, or the loss of arms q or p of the same chromosome (~50%). In fact, the high frequency of chromosome 9 losses strongly suggests that they are early events in TCC development. Deletions on 9p often involve the CDKN2A/ARF (P16) tumor suppressor region (9p21), in fact this abnormality is considered to be the most frequent event in TCC development. Interestingly, 9p deletions had been identified in normal appearing urothelium as well, confirming the hypothesis that genetic insults occur before a neoplasm becomes clinically visible. Further notable deletions involve 10q, which aberration perturbs the PTEN/MMAC1 tumor suppressor pathway, and 17p deletions which results in loss of TP53 function; the latter had been described to be correlated with tumor grading and staging. The cytogenetic heterogeneity of bladder cancer is well represented by the notion that while the loss of Y chromosome has been reported to be a sole abnormality, others described polysomy Y as an abnormality being correlated to tumor grading. Numerous other numerical chromosomal abnormalities have been described besides the above; the vast majority of these appear as chromosomal gains thus are presented as certain degree of polysomy. Such involved chromosomes are chromosomes 1,3,7,9,11,15,17 and 18

## The role of molecular cytogenetics in bladder cancer diagnostics

---

Due to conventional cytogenetics, comparative genomic hybridization (CGH) and DNA ploidy assays, the chromosomal instability of bladder cancer has been known for some time. Consequently, it has been proposed not once, that the problem of lacking a reliable, non-invasive diagnostic tool could be solved by cytogenetic analyses of exfoliated bladder cancer cells that can be found in voided urine. Conventional cytogenetic analysis is however not feasible for routine diagnostics, as cells from bladder TCC can be difficult to culture and overall are hard to karyotype. Accordingly, several investigations were carried out to develop a molecular cytogenetic method that would require no culturing and could be performed in interphase. Because there is no well defined sole cytogenetic abnormality, it was proposed, basically from the beginning, that the assay should aim to detect not one, but several different chromosomal abnormalities; i.e. to create a multiprobe FISH assay. After several attempts throughout nearly a decade, in 2000 a four probe FISH assay had been developed; somewhat later this assay was commercialized as Urovyison. The group having developed the probe-kit tested 10 frequently involved chromosomes; they analyzed 9 CEP probes (3,7,8,9,11,15,17,18 and Y) and one LSI probe (9p21) by creating three probe-groups. The sensitivities of the chromosomes were tested individually. The final group consisted of the 4 most efficient probes; three CEPs for visualizing chromosomes 3,7 and 17 with the colors of red, green and aqua, respectively, and 1 LSI for detecting p21 region of chromosome 9. This region, which is visualized by a 'gold' fluorochrome, contains the above mentioned p16 coding region, which – depending on alternative splicing – encodes several (at least three) proteins that function as tumor suppressors.

For the evaluation of the assay the recommended protocol is that a sample can be considered positive when containing  $\geq 4$  cells which are featured with polysomy of  $\geq 2$  CEPs and/or biallelic loss of 9p21 in  $\geq 12$  cells. Investigator should analyze 25 cells, however, if there is no positive result, the investigation continues until the result becomes positive; if needed, eventually by scanning the complete slide. An important momentum during evaluation is the targeting of the exfoliated bladder cancer cells. During analysis, cells are normally targeted according to DAPI (4,6-diamidine, 2-phenylindol dihydrochloride) morphology; by definition irregularly shaped cells that are featured with patchy, cloudy staining should be investigated. In 2001 the assay received the Food and Drug Administration's (FDA) approval for the follow up of patients with a history of bladder cancer; later in 2005 it received another FDA approval for the detection of bladder cancer in patients with microhematuria but no previous history of the disease. According to numerous studies the test has been shown to be superior to cytology in all grades. One of the most cited papers, reviewing results of 12 major publications describes sensitivity of cytology versus FISH in G 1-2-3 tumors as 18 vs. 50, 45 vs. 75% and 69 vs 90%, respectively. The same comparison for Ta-Tis-T1 and T2-4 tumors reveal sensitivities of 28 vs. 67%, 73 vs. 97%, 67 vs. 90% and 74 vs. 92%. The specificity of cytology however, appeared slightly better than

with FISH. Important to note, that both sensitivity and specificity moved on an extremely wide range; sensitivity varying between 8-89% and specificity between 29 and 90%. This variation is considered as yet another drawback of the method.

## Objectives

---

There are several well known weaknesses of multitarget fluorescence in situ hybridization performed on voided urine. One of these is the lack of objective definition of the target cells. Current target definition relies on DAPI morphology; however, voided urine is often compromised by degenerative changes, loaded with non-urothelial object (e.g. lymphocytes, squamous cells). Furthermore, phenotype does not necessary follow the genotypic aberration at the early stages of a tumor cell. Therefore, scanning of the samples at fluorescence light carries the same drawbacks as cytology, furthermore it is ambiguous and tiring.

Taking all this into consideration, we have set the aim of developing a method that enables the investigator to carry out FISH analysis strictly on previously identified uroepithelial cells.

1. During our preliminary studies we aimed to develop the most efficient consecutive phenol- and genotyping in control conditions. Our goal was

- to manage the highest possible level of automation
- to determine individual analytical accuracy of both phenotyping and consecutive genotyping
- to determine overall analytical accuracy of combined pheno-genotyping

2./a After developing targeted-FISH we had set the further aim of testing the method on clinical samples, thus to determine the methods diagnostic accuracy (i.e. sensitivity, specificity) and clinical utility.

2./b Another notable drawback of the Urovysion assay is that there is no gold standard for interpretation of results, neither there is a reliable and adaptable diagnostic cut-off. Current protocol provides binary (positive/negative) information, being based on a pre-defied absolute number cut-off, furthermore, currently there is no guidance on how to interpret ratio of positive cells (if one is to determine that at all). Some workgroups had already pointed out that a novel evaluation scheme that would take ratio of positive cells into consideration could have great diagnostic benefits.

In the light of these considerations, we have set the aim of testing an evaluation scheme different from current protocol and similar to approaches used in conjunction with other FISH assays. Our goal was to perform these investigations in a comparative manner; further goal

was to reveal association between tumor grade and stage, and the different evaluation approaches.

III. Amongst various advantages, the routine manual FISH analysis is featured with a number of well known drawbacks: (i) for statistical considerations often there is a need for investigating hundreds (if not thousands) of cells (ii) extremes of positivity or negativity could bias investigator and eventually lead to under- or overestimation of positive cell ratio (iii) interobserver or interlaboratory variability could notably hamper comparability of results. The probe set used in this study consist of 4 DNA probes; considering DAPI related targeting also, this indicate five decisions to be made in conjunction of every single cell, leaving the assay even more effected by the above drawbacks.

Consequently, being one of the first workgroups, we have set the aim of automating evaluation of this multiprobe FISH assay using a fully user trainable microscopic workstation. Our aim was to determine analytical accuracy, as well as diagnostic efficiency; further goal was to reveal any drawbacks, as well as to highlight technical improvement points.

## **Materials and Methods**

---

### **Samples**

---

Throughout our investigations we have performed experiments on both control and patient samples. Human peripheral blood and voided urine from healthy individuals served as negative controls; while a bladder cancer cell line (HT-1376) carrying multiple chromosomal aberrations served as positive control. Initially, samples from 73 patients were entered the study, out of which ten eventually had to be excluded due to technical problems. All patient samples were received from the Urology Department of the University of Pécs Medical School (UPMS), based on clinical suspicion following cystoscopy. Both histological and cytological diagnoses were given at the Department of Pathology (UPMS) by an experienced pathologist (NS). Histological classification, due to the above mentioned controversies, was given according to the 1973 WHO classification; cytological analysis of urine samples was based on international guidelines. All samples were prepared as cytological slide preparation.

The analytical accuracy (i.e. sensitivity, specificity) of each diagnostic approach detailed below was determined using the ‘true positive events + true negative events / all analyzed events’ formula; the value was also determined empirically, via correlation-data gained from dilution series created from positive and negative controls.

## Dual-mode scanning light microscope (SLM)

---

For all our cytometric investigations we used a dual-mode light microscopy based workstation. The microscope (Zeiss Axioplan-MOT II.) is able to operate using both transmitted and fluorescence light. It is equipped with a motorized, eight slide Merzhäuser scanning stage, a BW CCD camera a motorized filter wheel and a 100W mercury lamp. The whole instrument is connected to a computer running several special softwares such as Metacyte-Metafer or Isis (Metasystems, Germany). These softwares enable automated analysis including complete automation of scanning stage movements to the  $\mu\text{m}$  precision, and filter changing; it also enables image capturing and digital image processing, furthermore, complex cytometric measurements. One of the most valuable feature of the system is that it records exact 3D coordinate of captured objects; accordingly, these are easily relocatable latter.

## Semi-automated phenotyping (targeting)

---

For phenotyping cells of urothelial origin we applied a monoclonal antibody against cytokeratin 7. The reason why we choose CK7 was because this low molecular weight keratin (LMWK) is strongly and consistently expressed in all layers of the urothelium, as well as in all neoplasms that arise from these cell types. On the other hand the marker is not expressed by other cell types that can be found in urine, such as inflammatory cells or squamous cells.

For phenotyping we performed standard immunocytochemistry on slides; the antibody was visualized using secondary antibodies binding to a dextran ‘backbone’ and 3-Amino-9-Ethylcarbazole. The CK7 and hematoxylin stained samples were scanned automatically at transmitted light on a manually defined area of the slide. The system identified the cells according to pre-defined morphometric parameters, these were: counterstain gray level threshold (12%), minimum object area ( $20\mu\text{m}^2$ ), maximum object area – 1000 ( $\mu\text{m}^2$ ) maximum concavity depth (1,00) maximum aspect ratio (2,905). Following scanning the CK7 positive elements were separated according to their immunostain related pixel intensities. The process was based on simultaneous gating on the histograms depicting staining intensity related features (total intensity within the contour of the nucleus – TIC; mean intensity within the contour of the nucleus – MIIc). The selection thresholds were determined based on control measurements on 28,000 negative and 71,000 positive cells. In brief, we measured the staining intensities of all control cells and investigated what portion of the cells would still be ‘visible’ at various intensity thresholds; in case of positive controls this determined the true positivity (sensitivity) at various staining intensities, whereas by negative data from negative controls, it depicted false positives. These values were portrayed simultaneously on a coordinate plane; this way we could analyze false positivity and negativity at a given staining intensity value (arbitrary staining unit – asu) simultaneously. As we used two intensity related

features eventually we had to determine threshold (and thus specificity and sensitivity) using two coordinate planes. After determining the biologically (and not just mathematically) optimal cut-offs, following analysis was carried out by gating on histograms delineating 'asu's for M1c and T1c.

## Genotyping of phenotypically preselected cells; comparative study

---

Just like in case of FISH-alone analysis following immunocytochemistry slides were thoroughly washed in PBS (3x'5), following which slides were treated with Carnoy fixative (methanol: glacial acetic acid = 3:1, -20°C) overnight. Following sufficient air drying the multiprobe FISH assay (Urovysion) was performed according to manufacturer's instructions.

### Manual evaluation

---

The comparative FISH investigations implied three main steps: (1) Locating target cells: when applying Urovysion-alone, targeting was based on DAPI morphology, and carried out manually however when applying targeted-Urovysion it was based on CK7 positivity, using automated object relocation. (2) Evaluation of signal pattern: positive cell definition was based on manufacturer's instructions regarding both approaches. That is, a cell was considered positive if gains for at least two different chromosomes in the same cell and/or bi-allelic loss of 9p21 could be detected. (3) Interpretation of FISH results: results of both Urovysion-alone and targeted-Urovysion were evaluated by both conventional and statistical method. Conventional evaluation relied on manufacturer's recommendation (absolute diagnostic cut-off values: a sample is positive when it contains: 4 cells with gains for at least two different chromosomes and/or 12 cells with bi-allelic loss of 9p21), while statistical approach used calculated thresholds. Cut-offs were derived using the FP + 2SD rule and were set to 10.3% for Urovysion-alone and 1% for targeted-Urovysion.

As we hypothesized that the quantity of exfoliated cells could carry diagnostic information in itself, the required sample size (number of cells to be analyzed) were also calculated to gain sufficient (95%) statistical power (confidence limit) at the detected positivity rates, for the determined analytical sensitivities. The values were calculated on a hypothetical scale, showing positivity rates from 10—90%, using the related analytical sensitivities. The applied formula:  $N = [Z_{\alpha}(p_0(1-p_0))^{-1/2} Z_{\beta}(p_a(1-p_a))^{-1/2} / p_0 - p_a]^2$  was based on a one-sided, hypothesis test for a binominal proportion, and was previously detailed by Dewald et al. (currently, it is used by the American College of Medical Genetics as a guideline in "Standards and Guidelines for Clinical Genetics"). A sample was considered positive if it met the criteria of sufficient cellularity and its positivity rate exceeded determined diagnostic cut-off (double statistical criteria). The prior criterion, in fact, would rather be sufficient as an exclusion condition, however as we predicted that the amount of exfoliated urothelial cells

might represent diagnostically important information, this measure was also treated as a diagnostic criterion; if we could not determine a samples positive status  $\geq 95\%$  the sample was considered negative.

## Automated Evaluation

---

### *Capture and Focusing*

---

After setting both the search window and reference object manually, two initial automated focusing steps were executed: grid focusing and fine focusing. For the former, an autofocus algorithm was applied at X/Y positions on a virtual regular square grid, layered over the search window. X/Y distances of the focus positions were set to 1000  $\mu\text{m}$  and autofocus was performed using 9 focus planes with a Z spacing variation of 1.75  $\mu\text{m}$ . Once the optimum focus was established for all grid autofocus points, it was interpolated for the center of every image field of the search window. Consecutive fine focusing implied 7 focus planes for which 1  $\mu\text{m}$  Z spacing adjustment was performed before each image field capturing. Overall detection of nuclei was enabled in a vertical range of  $\sim 19$   $\mu\text{m}$ .

Digital image fields were captured using a 40x objective. Exposure control was automatic, with a maximum integration time of 1.00 s for the DAPI channel and 2.00 s for the signal channels. Maximum saturation area was set to 4  $\mu\text{m}^2$  and 1  $\mu\text{m}^2$  in DAPI and signal channels, respectively. In the DAPI channel, images were captured in a single plane, as opposed to the signal channels where 5 focus planes were used, with Z steps of 0.5  $\mu\text{m}$ . Internal camera gain factor was set to a relatively low value of 2.0 to reduce electronic noise.

### *Training*

---

The training process started by performing an automatic scan using a standard slide (made from urine of a patient with confirmed diagnosis of malignancy) to create digital training fields. The following process mainly consisted of 3 related steps: classification, testing and optimizing.

Classification started by manually classifying nuclei in the DAPI channel. This classification was used as the reference during consecutive tests of automated cell selection. The cell selection parameters were as follows: counterstain object threshold, minimum object area, maximum object area, maximum concavity depth and aspect ratio. Parameter adjustments and repeated testing were performed until performance was optimal. Classification of FISH signals of the same images were performed in all four channels separately. For spot counting, the fundamental features in all channels were: spot measurement area, minimum spot distance and minimum spot relative intensity. During automated test runs (channel by channel), the system considered the previously classified objects and their signal patterns as an absolute reference. After analyzing false hits,

recognition parameters were adjusted accordingly and the test was repeated. This process was iterated multiple times to optimize various parameters (including image and cell processing operations) and to achieve optimal accuracy for each signal separately.

When individual detection performance was optimal, a comprehensive test run was performed by re-scanning the same digital slide. This way a comprehensive testing of automated analysis could have been carried out via simulating the likely procedure for routine slide scanning. Individual nuclei and their signal pattern were displayed in the gallery and overall detection accuracy was verified interactively.

## Results and Conclusions

### Semi-automated pheno- and genotyping of uroepithelial cells using chromogenic immunocytochemistry and multiprobe fluorescence in situ hybridization

Comparing specificities, Urovysion-alone with conventional enumeration showed the poorest performance, as it appeared even weaker than cytology (Table 1.). The false positive cases were the ones where ambiguous DAPI morphology indicated extensive screening. This is bound to happen as proportional to analytic FP a certain percentage (of any cell population) will certainly appear positive. Our results showed that by applying a cut-off that relies on locally determined analytic FP (FP+2SD) this phenomenon can be avoided, resulting 100% specificity.

Table 1. Comparison of diagnostic efficiency values by grades, regarding different approaches: cytology, Urovysion-alone and targeted-Urovysion

	Cytology	FISH					
		Urovysion-alone			targeted-Urovysion		
		Conventional Evaluation (1/a)	Statistical Evaluation (1/b)	mean of 1/a,b	Conventional Evaluation (2/a)	Statistical Evaluation (2/b)	mean of 2/a,b
<b>Specificity</b>	86% (6/7)	71% (5/7)	100% (7/7)	86%	100% (7/7)	100% (7/7)	100%
<b>Sensitivity<sub>overall</sub></b>	60% (21/35)	85% (28/33)	76% (25/33)	80%	97% (34/35)	89% (31/35)	93%
G1	67%	100%	100%	100%	100%	100%	100%
G2	50%	78%	61%	69%	95%	85%	90%
G3	75%	92%	92%	92%	100%	92%	96%
<b>Accuracy</b>	64% (27/42)	82% (33/40)	80% 32/40)	81%	98% (41/42)	91% (38/42)	94%

Targeted-Urovysion in general performed a steady 100% specificity as well. Interestingly, for both conventional and statistical evaluation, the low number of urothelial cells was the eventual cause of not exceeding absolute cut-off value and failing to meet the criterion of minimum cellularity, respectively. Considering also that in histologically positive cases only a small percentage failed to meet this criterion, this suggests that low urothelial (CK7 positive) cellularity alone might have a notable negative predictive value.

Regarding overall sensitivity, cytology performed weaker than any interpretation of FISH. In general, targeted-Urovysion was more sensitive than Urovysion-alone, especially in low-grade and low stage lesions. In fact, as Ta stage tumors are the most hardly diagnosable ones using the conventional UroVysion-alone evaluation, the prominent increase of sensitivity in this early stage category is considered to be one of the biggest benefits of targeted-Urovysion (Table 2.).

Table 2. Comparing diagnostic sensitivities experienced in different pathologic stages

	Cytology	FISH			
		Urovysion-alone		targeted-Urovysion	
		Conventional Evaluation (1/a)	Statistical Evaluation (1/b)	Conventional Evaluation (2/a)	Statistical Evaluation (2/b)
<b>pTa</b>	27% (4/15)	71% (10/14)	57% (8/14)	93% (14/15)	87% (13/15)
<b>T1</b>	71% (10/14)	85% (11/13)	77% (10/13)	100% (14/14)	86% (12/14)
<b>mean (pTa, T1)</b>	48%	72%		91%	
<b>T2</b>	83% (5/6)	100% (6/6)	100% (6/6)	100% (6/6)	100% (6/6)

The false negative cases of statistically evaluated Urovysion-alone came from failing to exceed diagnostic cut-off. This contributes to the fact that the lower analytic specificity (=higher analytic FP) of Urovysion-alone determines higher diagnostic cut-off (which is determined as: analytic FP + 2SD), furthermore, ambiguous DAPI morphology could just as well cause overestimation of the negative cell ratio.

Taking all results of objective targeting into account, it shows that increased analytical accuracy gained by restricting analyzed population to preselected urothelial cells, manifests in increased diagnostic accuracy as well. Furthermore, the approach requires less cytopathological expertise and time, since investigator can objectively screen the whole slide without having to analyze every single cell. Additionally, in the light of recent publications<sup>17</sup>, a further important aspect of immunophenotyping prior to urine FISH analysis should be considered; that is, the potential of clarifying the origin of tumorous signal pattern depicting

nuclei in ambiguous cases or when non-urothelial (e.g. cervical) carcinomas are simultaneously present with bladder carcinoma.

When comparing diagnostic accuracy from the aspect of evaluation forms, the statistical approach appeared somewhat less accurate, due to drop of sensitivity. This suggests that our statistical criteria might have been too stringent and need some further modification, possibly regarding the criterion of sufficient cellularity. However, the determination of ratio of positivity provided important additional quantitative information, as results of both targeted and non-targeted approaches were associated with tumor grading and staging as well. In fact, targeted analysis showed even tighter association and differentiated more explicitly especially high stage, high grade tumors from other malignancies. Differences were statistically non-significant with non-targeted approach (G1-2-3, 12,9%, 28,7%, 33,7%, respectively;  $p=0,1097$ ), but significant with targeted-approach (G1-2-3, 30,87%, 42,5%, 84,4%, respectively;  $p=0,0019$ ). Again proving that by objectively excluding non-urothelial cells, more definite information could be gained regarding tumor load. The potential benefit of percentage based evaluation has been highlighted previously. Moreover, it had been reported that the percentage of abnormal cells predicts tumor recurrence and progression. It indeed should be considered, that a qualitative diagnosis (positive or negative) in conjunction with a quantitative molecular test bereaves that method from having at least the chance of providing with further classification and therapy relevant information. However, it also should be considered that FISH assays have great inter-laboratory variability. A related report highlights: different fractions of contaminated cells counted, and differences in analytical accuracy or hybridization efficiency, as potential causes. Because of its multitarget feature, these factors pertain to Urovysion even more. Accordingly, the adoption of the same preset absolute value cut-off in different diagnostic laboratories may fail to produce consistent, comparable results. A fixed, seemingly standardized cut-off might just in fact be a feature that hampers standardization in lack of which comparison between single studies is difficult. Consequently, we agree that percentage based evaluation could provide valuable information, however the application of locally set statistically derived cut-offs should also be considered.

In conclusion, immunophenotypic targeting simplifies Urovysion's evaluation procedure, increases diagnostic accuracy, and it could play an important role in differential diagnosis. Determination of positive cell ratio could provide with the potential of molecular classification and increase the therapy relevant value of urine based FISH analysis. The application of statistically derived cut-offs is feasible and if laboratories would set these values individually but according to a harmonized, consensus protocol – similar to other FISH assays - , variances in results might decrease. This would facilitate interlaboratory comparison and might be an important step towards the long wished standardization of Urovysion.

## Full automation of signal pattern evaluation of multitarget FISH probe-set, containing both CEP and LSI probes

---

Urovysion is a promising non-invasive technique suitable for replacing or at least complementing current less sensitive or invasive standard methods for detecting bladder cancer. However, enumeration of signal pattern is laborious and requires considerable experience in FISH evaluation. In addition, recent studies have highlighted the potential benefits of percentage based (quantitative) evaluation, as opposed to recommended (qualitative) modus, relying on absolute numbers. A quantitative method of such nature will further increase the number of analyzed cells. Definitely, the automation of the signal enumeration process would be of great benefit as it has been for other FISH assays. Numerous reports have detailed their attempts in automating enumeration of FISH signals specific for numerical or structural chromosomal abnormalities using a variety of systems. However, reports discussing automation of multitarget FISH assays are limited, and to our knowledge, an in-depth report on the process of automating Urovysion's signal enumeration with a user-trainable system has not been published before. Thus, in the current study our objective was to investigate this process in detail to determine efficiency and feasibility and to reveal its potential pitfalls.

During parameter adjustment of basic object (nuclei) recognition, one of our main goals was to achieve high accuracy, primarily via high sensitivity rather than high specificity. This is important, because false positives like non-cellular elements would obviously be disqualified in the consecutive signal pattern analysis, thus would not affect final accuracy. Undetected cellular objects, however, would decrease overall analytic and diagnostic sensitivity. Overall, cell selection was performed with a false positivity of 7,3%, and a false negativity rate of 3,4%. On the other hand, when adjusting for signal recognition performance, there was no preference for sensitivity or specificity. Similar parameter values were sufficient for CEPs, but LSI 9p21 required notably dissimilar settings, mainly because this locus specific probe creates smaller, dimmer signals compared to CEP ones. It is particularly important because its bi-allelic deletion is by itself a tumor indicating pattern and is known to be one of the initial genetic changes in bladder cancer. As the gold signals are weak, automatic exposure implied a 2 to 3 times longer integration time, resulting in an intensive, inhomogeneous background, part of which would appear as signals. Higher relative intensity and smaller spot area criteria had to be determined, to have signals excel from background, and gain adequate detection accuracy.

Even though the average spot enumeration accuracy for different probes was above 80% (accuracy for CEPs 3, 7 and 17 was 77,3, 82,6 and 81%, while for LSI 9p21 it was 87,1%), the overall spot enumeration accuracy was ~54%; which was slightly higher than suspected (~45%). This confirms that the errors occurring in different channels are mostly independent, therefore generate an overall error rate resembling the value of their mathematic

multiplication. Despite this low value, cell classification accuracy was high (both analytic specificity and sensitivity was >90%). This can be explained by the complex criteria used by Urovysion for defining a positive cell, whereby the eventual classification of a cell is not necessarily affected by an error in one signal channel. However, the parameter that does influence cell classification is the nature of spot misclassifications. In most cases misclassification of CEPs was moderate, in the region of missing true spot counts by 1 or 2 values (i.e. 5 spots instead of true 6 spots). In this cohort, the decisive effect of misclassification was suspected to decrease with the increased degree of polysomy, since according to the manufacturer's recommendation (currently) there is no diagnostic difference between 5 and 6 spots, as both values indicate chromosomal gain, thereby implicating tumor. Differences in analytic specificity and sensitivity (92.2% and 97.6%, respectively) confirmed this assumption. The automated analysis showed a decline in the efficiency of determining positive/negative status in negative (diploid) samples, compared to highly polysomic positive cell population, where the effect of moderate misclassification was much less. This phenomenon was also the reason why more than twice the number of negative cells had to be disqualified than positive cells, due to zero detected signals for one or more CEPs (5,21% ± 1,78 vs. 12,07% ± 2,94%). When modeling actual conditions with dilution series, results of automated analysis tightly correlated with the expected values ( $r^2=0,98$ ). However, we experienced some underestimation of the negative cell numbers, possibly due to exclusions made for zero CEP signal. This mostly contributed to the 'degree of polysomy'-dependent misclassification as emphasized above, and also to occasional differences in the signal intensities between negative and positive control cells.

According to our clinical data, diagnostic sensitivity increased with the stage and grade, applying either manual or automated methods. These findings are in accordance with the literature. However, the differences were more pronounced with automated analysis, which is suspected to contribute to two factors. First, with increase in both frequency of polysomic cells and degree of polysomy, the effect of moderate misclassification of the system gradually diminished. This further confirms the 'degree of polysomy'-dependent cell misclassification associated with automated analysis. Consequently, gaining a high level of automation, where human factors could be totally eliminated from the method of signal pattern enumeration, would be a considerably easier task to accomplish in cases with high polysomy (e.g high stage/grade cases). It is in fact questionable whether automated cell classification with Urovysion (or any assay, aiming to detect multiple numerical chromosomal changes) would ever reach the accuracy level of human investigation in cases with low polysomy. Second, in some cases, the dilution effect of negative results of non-urothelial nuclei resulted in a fallacious low positivity, preventing those cases from exceeding diagnostic positivity cutoff. Although, this phenomenon had a much less grade/stage preference, it still represents an important issue. The problem of DAPI morphology being ambiguous and providing limited targeting information has been

underlined earlier, in relation with manual analysis. Since a non-interactive, fully automated system could identify objects based only on well defined objective criteria, it could be even more limited than the human eye in such a subjective recognition task. In our previous reports we have described potential benefits of automated, objective (immunophenotypic) targeting and proved that it increases accuracy of consecutive, manually evaluated Urovysion assay. The combination of that approach with the current automated spot counting method would be of even greater significance. Moreover, it would mean that a comprehensively and fully automated diagnostic method identifies and analyses urothelial tumor cells solely by objective phenotypic and genotypic criteria, minimizing the human factor.

## Summary

---

During our investigations we were the first workgroup to develop an automated image-cytometric approach for the consecutive pheno- and genotyping of uroepithelial cells in voided urine; using chromogenic immunocytochemistry for the prior and fluorescence in situ hybridization for the latter. As a result, we

- determined the detailed technical specifications of the semi-automated chromogenic preselection; these data could serve as the basis of any similar chromogenic immunophenotypic preselection methods.
- concluded that the combined method simplifies FISH analysis on voided urine, moreover it objectifies initial targeting which in this sense does not require notable cytological experience.
- revealed that the combined method is featured with enhanced analytical accuracy compared to FISH-alone.

Performing comparative studies, we determined the combined methods diagnostic efficiency, during which, apart from comparing 'targeted-FISH' to 'FISH-alone', we have also compared the evaluation design of applying a locally set diagnostic cut-off and considering positive cell ratio opposed to recommended evaluation scheme. According to our results, we

- concluded that the combined method's increased analytical efficiency manifests in increased diagnostic efficiency and therefore enhanced clinical utility as well.
- concluded that one of the biggest advantages of targeted approach is that it is featured with increased sensitivity in the diagnostically challenging low grade, low stage tumors. The notion is due to the comprehensive analysis of the whole sample as well as to the low diagnostic cut-off, being caused by high analytical

specificity. These features could enable the method to stand as an effective clinical monitoring approach.

- confirmed the hypothesis that the ratio of FISH positive cells correlates with tumor grading and staging; in fact, we described that only targeted-genotyping could give statistically significant results in this regard. Conclusively, our method provides a more realistic view on the true tumor load.

We were the first ones to automate evaluation of multiprobe FISH assay on voided urine using a user trainable workstation. We determined the method's analytical accuracy as well as detailed specifications mandatory for automation. The latter could serve as a basis for other automation of evaluation of multiprobe assays aiming to investigate complex numerical aberrations. According to our results, we

- confirmed the hypothesis that the errors occurring in different channels are mostly independent, thus generating overall error rates resembling the value of their mathematic multiplication. Furthermore, because of the notion it can be concluded that the automated signal recognition efficiency of multiprobe assays, depends on the positive cell definition just as much as it does on the individual signal recognition accuracy of different channels. Since positive cell definition is most notably effected by the genetic complexity (i.e. instability) of the given entity, we can conclude that the individual signal recognition efficiency experienced with multiprobe assays aiming to detect genetically instable tumors are less important than in cases of single or dual probe assays aiming to detect an aberration that is tumor indicating by itself.
- concluded, that in case that the individual misclassifications are moderate, the effect of miss-hits on single cell classification decrease with the increased degree of polysomy (i.e. grade and stage).

## **Acknowledgements**

---

I would like to express my gratitude to Professor Béla Melegh, under whose supervision I have been working since 2008. I am also grateful to my father Professor László Pajor who was the fundamental motivational force of my initial interest in natural sciences and who latter greatly affected my approach as a researcher. I am obliged to Dr Norbert Süle, who has helped my work with useful suggestions and as a pathologist, always reminded me of the importance of diagnostic utility. I am grateful to Dr Donát Alpár and Dr Béla Kajtár for their very useful suggestions and for the countless scientific discussions, and to Marika Kneif who

introduced me into laboratory life, and who was always ready to help me in case of any technical problems.

Besides the above, I would like to express my acknowledgements to Vera Kalász for her helping and teaching attitude in cell culturing; to urologists Dr László Somogyi and Professor László Farkas whose help and cooperation was indispensable for investigating diagnostic utility; to Dr Pál Jáksó, who was so helpful in preparing digital images for publications and presentations, and to other members of the Pathology Department who helped me along the way.

Last but not least, I am very grateful to my family for all their support; I am thankful for my wife's patients, and I owe special gratitude to my children, Gréta and Berci who, without even realizing, helped me through the hard days, while I was "...peeking at those colorful spots..."

## List of publications

---

### Publications related to the thesis

---

#### Original publications

---

**Pajor G**, Alpar D, Kajtar B, Melegh B, Somogyi L, Kneif M, Bollmann D, Pajor L, Sule N. Automated signal pattern evaluation of a bladder cancer specific multiprobe-FISH assay applying a user-trainable workstation. *Microscopy Research and Technique*, Accepted for publication (2011.11.07.). *IF.:* 1,712

**Pajor G**, Somogyi L, Melegh B, Alpar D, Kajtar B, Farkas L, Kneif M, Bollmann D, Pajor L, Sule N. Urovysion: Considerations on modifying current evaluation scheme, including immunophenotypic targeting and locally set, statistically derived diagnostic criteria. *Cytometry Part A* 2011;79(5):375-82. *IF.:* 3,753

**Pajor G**, Süle N, Alpár D, Kajtár B, Kneif M, Bollmann D, Somogyi L, Pajor L. Increased efficiency of detecting genetically aberrant cells by UroVysion test on voided urine specimens using automated immunophenotypical pre-selection of uroepithelial cells. *Cytometry Part A* 2008;73(3):259-65. *IF.:* 3,259

Citable abstracts

---

**Pajor G**, Kneif M, Csala J, Farkas L, Pajor L, Süle N. Detection of primary urothelial carcinoma in voided urine specimen using four probe FISH assay combined with automated microscopic system, a prospective study. *Virchows Archiv* 2005;447(2). *IF.:* 2,227

---

***Impact factor: 10,951***

Other publications

---

**Pajor G**, Kajtár B, Pajor L, Alpár D. State-of-the-art FISHing: automated analysis of cytogenetic aberrations in interphase nuclei. REVIEW. *Cytometry Part A*, *accepted for publication* 2012.05.22. *IF.:* 3,753

Bollmann D, Bollmann M, Bankfalvi A, Heller H, Bollmann R, **Pajor G**, Hildenbrand R. Quantitative molecular grading of bladder tumours: a tool for objective assessment of the biological potential of urothelial neoplasias. *Oncol Rep* 2009;(1):39-47. *IF.:* 1,588

Alpar D, Hermes J, Poto L, Laszlo R, Kereskai L, Jakso P, **Pajor G**, Pajor L, Kajtar B. Automated FISH Analysis Using Dual-Fusion and Break-Apart Probes on Paraffin-Embedded Tissue Sections. *Cytometry Part A* 2008;73A:651-657. *IF.:* 3,259

Buzogány I, Bagheri F, Süle N, Magyarlaki T, Kalmár-Nagy K, Farkas L and **Pajor G**. Association between Carcinoma and the Transplanted Kidney. *Anticancer Research* 2006;26: 751-754. *IF.:* 1,479

---

***Impact factor: 10,079***

---

***Cumulative impact factor: 21,030***