Novel urinary biomarkers for the detection of bladder cancer: A systematic review

Wei Shen Tan\textsuperscript{a,b}, Wei Phin Tan\textsuperscript{c}, Mae-Yen Tan\textsuperscript{d}, Pramit Khetrapal\textsuperscript{a,b}, Liqin Dong\textsuperscript{e}, Patricia deWinter\textsuperscript{a}, Andrew Feber\textsuperscript{e}, John D Kelly\textsuperscript{a,b}

\textsuperscript{a} Division of Surgery and Interventional Science, University College London, 3rd floor Charles Bell House, 43-45 Foley Street, London W1W 7TS, UK.
\textsuperscript{b} Department of Urology, University College London Hospital at Westmoreland Street, 16-18 Westmoreland Street, London W1G 8PH, UK.
\textsuperscript{c} Department of Urology, Rush University Medical Center, 1653 W Congress Pkwy, Chicago, IL 60612, USA.
\textsuperscript{d} School of Public Health, London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT, UK.
\textsuperscript{e} UCL Cancer Institute, University College London, Paul O’Gorman Building, 72 Huntley Street, London WC1E 6DD, UK.

*Corresponding author:
Wei Shen Tan
Division of Surgery & Interventional Science,
University College London,
3rd floor Charles Bell House
43-45 Foley Street
London W1W 7TS
UK
Tel: +44(0)20 7679 6490
Fax: +44(0)20 7679 6470
Email: wei.tan@ucl.ac.uk

Funding support: No specific funding to disclose

Conflict of interest disclosure: No conflict of interest to disclose
Novel urinary biomarkers for the detection of bladder cancer: A systematic review
Abstract

Background
Urinary biomarkers for the diagnosis of bladder cancer represents an area of considerable research which has been tested in both patients presenting with haematuria and non-muscle invasive bladder cancer patients requiring surveillance cystoscopy. In this systematic review, we identify and appraise the diagnostic sensitive and specificity of reported novel biomarkers of different 'omic' class and highlight promising biomarkers investigated to date.

Methods
A MEDLINE/ Pubmed systematic search was performed between January 2013 and July 2017 using the following keywords: (bladder cancer OR transitional cell carcinoma OR urothelial cell carcinoma) AND (detection OR diagnosis) AND urine AND (biomarker OR assay). All studies had a minimum of 20 patients in both bladder cancer and control arms and reported sensitivity and/ or specificity and/ or receiver operating characteristics (ROC) curve. QUADAS-2 tool was used to assess risk of bias and applicability of studies. The search protocol was registered in the PROSPERO database (CRD42016049918).

Results
Systematic search yielded 115 reports were included for analysis. In single target biomarkers had a sensitivity of 2-94%, specificity of 46-100%, positive predictive value (PPV) of 47-100% and negative predictive value (NPV) of 21-94%. Multi-target biomarkers achieved a sensitivity of 24-100%, specificity of 48-100%, PPV of 42-95% and NPV of 32-100%. 50 studies achieved a sensitivity and specificity of ≥ 80%. Protein (n=59) and transcriptomic (n=21) biomarkers represents the most studied biomarkers. Multi-target biomarker panels had a better diagnostic accuracy compared to single biomarker targets. Urinary cytology with urinary biomarkers improved the diagnostic ability of the biomarker. The sensitivity and specificity of biomarkers were higher for primary diagnosis compared to patients in the surveillance setting. Most studies were case control studies and did not have a
predefined threshold to determine a positive test result indicating a possible risk of bias.

Conclusion
This comprehensive systematic review provides an update on urinary biomarkers of different ‘omic’ class and highlights promising biomarkers. Few biomarkers achieve a high sensitivity and negative predictive value. Such biomarkers will require external validation in a prospective observational setting before adoption in clinical practice.

**Keywords:** Bladder cancer; Biomarker, Diagnosis, Systematic review, Urine

**Highlights:**
- Multi-target biomarker panels had a better diagnostic accuracy compared to single biomarker targets
- The sensitivity and specificity of biomarkers were higher for primary diagnosis compared to patients in the surveillance setting
- Most studies were case control studies and did not have a predefined threshold to determine a positive test result indicating a possible risk of bias
- Prospectively field tested to validate biomarkers for the detection of bladder cancer are required
- Utilization of next generation sequencing with machine learning represents a promising approach for biomarker discovery
Introduction

Bladder cancer is the eight most common cancer and ranks 13th in terms of cancer associated mortality\(^1\). Haematuria, a cardinal symptom for bladder cancer, has a positive predictive value of 8% and this rises to as high as 18.7% in men ≥ 70 years\(^2\). Patients presenting with haematuria undergo investigations including cystoscopy and upper tract imaging. Eighty percent of patients with bladder cancer have non-muscle invasive bladder cancer (NMIBC) at presentation. While this is favorable compared to muscle invasive bladder cancer (MIBC), up to 50% of NMIBC cases recur and 20% will progress within 5 years\(^3\). Due to this high recurrence rate, regular surveillance cystoscopy is recommended, and the surveillance interval can be as frequent as three monthly in high risk disease\(^4\).

Cystoscopy remains the gold standard for the detection of bladder cancer in patients investigated following haematuria and in patients requiring surveillance for recurrent disease following resection of the initial tumour. However, it is not without morbidity and up to 5.5% of patients may develop a urinary tract infection\(^5\). The requirement for life long surveillance in high risk patients have significant healthcare cost implications. Hence, there is an urgent need to develop a highly specific and sensitive urinary biomarker for the detection of bladder cancer.

Currently the US Food and Drug Administration has approved six urinary assays for clinical use; BTA stat (Polymedco), BTA TRAK (Polymedco), NMP22 (Matritech), NMP22 BladderCheck Test (Alere), uCyt (Scimedx) and UroVysion (Abbott Molecular). The tests perform with overall sensitivity between 57-82% and specificity between 74-88%\(^6\). Although sensitivity is higher in high grade and stage tumours, cystoscopy remains the gold standard for detection of bladder cancer, with a sensitivity as high as 98%\(^7\). Thus, none of these assays are approved to be used without cystoscopy.
There has been considerable interest in the development in urinary biomarkers as evident by the large number of published reports. While many show promising results, few have been reproduced in subsequent independent validation studies. Traditional assays have been designed for single targets or small panel assays restrained by the technology and assay performance. More recently, next generation sequencing and advancements in bioinformatics has enabled a paradigm shift whereby biomarker panels comprise multiple targets has been utilised using small quantities of input DNA.

In this systematic review, a literature search between January 2013 to July 2017 was performed to provide an update of urinary biomarkers for the detection of bladder cancer across the spectrum of protein, genomic, epigenetic and transcriptomic biomarkers. The purpose of this study is to highlight promising biomarkers which may have clinical utility in the future.
Methods

Literature search
A systematic search of the literature was performed using MEDLINE/PubMed to identify articles evaluating novel urine biomarkers for the detection of bladder cancer. A comprehensive literature search was performed between 1st January 2013 and 31st July 2017 using the following keywords and MeSH terms: (bladder cancer OR transitional cell carcinoma OR urothelial cell carcinoma) AND (detection OR diagnosis) AND urine AND (biomarker OR assay). The search protocol was registered in the PROSPERO database (CRD42016049918).

Study selection
Article selected were written in English and reported the diagnostic characteristics of novel urinary biomarkers for the detection of bladder cancer. Following screening of abstracts to exclude review articles, comments and letters to the editor or non-relevant articles, each manuscript was reviewed and data was extracted and its references searched for relevant missing manuscripts.

All studies required a minimum of $\geq 20$ patients in both bladder cancer and control arm to be included and report both sensitivity and/ or specificity and/ or receiver operating characteristics (ROC) curve. The presence of bladder cancer was defined as the presence of cancer at histopathological examination following transurethral resection of bladder cancer. Biomarkers were classified to protein, genomic, epigenetic, transcriptomic and combination of different ‘omic’ biomarkers.
All abstracts and full text were independently screened by two investigators. Where there were disagreements, this was discussed with a third investigator and resolved by a consensus view. Cohort and cross-sectional studies were included.

**Data extraction and quality assessment**

Data was extracted from selected studies about type and biomarker used, assay used, study design, percentage of low grade cancer assayed, urine collection details and number of patients with bladder cancer and controls (WST, WPT, MYT, PK). Where more than one patient cohort were described, the final validation patient group was used. Low grade tumours were defined according to EAU risk classification\(^8\). A 2 X 2 table with number of true-positive, false-positive, true-negative, and false-negative results from published sample sizes was constructed to determine the sensitivity, specificity, positive (PPV) and negative predictive value (NPV) where available. ROC curve where reported was included. A second investigator confirmed data were extracted accurately. QUADAS-2 tool was used to assess risk of bias and concerns about applicability of studies\(^9\).
Results
Characterization of studies
The PRISMA flowchart is shown in Figure 1. The database search identified 646 articles and after the addition of other relevant articles, a total of 656 abstracts were screened. Dual review of abstracts and titles excluded 377 studies which were not original research, not in English or unrelated articles. A further 164 studies were excluded after full text review as they did not meet the inclusion criteria leaving 115 articles which were included for analysis.

Articles were then classified to the following biomarkers: protein (n=59), genomic (n=7), epigenetic (n=19), transcriptomic (n=21) or combination of different ‘omic’ biomarkers (n=10). Twenty five protein, 1 genomic, 8 epigenetic, 10 transcriptomic and 6 combination of different ‘omic’ biomarkers had a sensitivity and specificity ≥ 80%. Studies with a sensitivity and specificity of < 80% are shown in the Appendix A2-A6).

Of the studies with a sensitivity and specificity ≥ 80%, most of these studies were designed as case control with selected groups comprising of urine from bladder cancer and control cases indicating selection bias (Appendix A1). Four prospective observational studies with some incorporating sequential urine sampling with surveillance cystoscopy although none had pre-planned statistical power calculations. Twenty three studies had a low risk of bias in determining the characteristics of the index test according to the QUADAS-2 tool.
Protein biomarkers

Protein based biomarkers were the most commonly tested biomarker for the detection of bladder cancer and used either immunoassays (n=35) or spectrometry (n=9) for protein quantification. Multiple protein targets were tested in 14 studies using multiplex immunoassay platforms interrogating between 3-10 biomarkers (Table 1 & A2).

Fourteen tests which tested an individual protein biomarker reporting a sensitivity and specificity ≥ 80%,10-18, 20, 21, 27, 30, 34 (Table 1). Of these, Orosomucoid 1 (ORM1), an acute phase transport protein, identified using mass spectrometry was quantified using ELISA of urine with a sensitivity of 92%, specificity of 94% and an ROC of 0.965. A separate study of 152 patients reported good diagnostic accuracy using the serine protease, HtrA1, and achieved a sensitivity of 93% and specificity of 96%.

Survivin is a protein which is implicated in the inhibition of apoptosis, has been investigated by a number of studies12, 13, 60. Quantification of survivin using ELISA reports a sensitivity of 71-85% with a specificity of 81-95%. Soluble Fas was reported by two studies and showed varying sensitivity of 51% and 88% which suggesting a lack of reproducibility16, 61.

Amplified in breast cancer 1 (AIB1) which has been shown to promote cell proliferation via AKT pathway had a sensitivity and specificity of 80% and 86% respectively. When combined with eukaryotic initiation factor 2 (EIF5A2) and nuclear matrix protein (NMP22) this increased to a sensitivity of 89%, specificity of 91% and ROC of 0.898. Other reports on single protein biomarkers include apurinic/apyrimidinic endonuclease 1/redox factor-1 (APE1/Ref-1), apolipoprotein A-I (Apo-A1), calprotectin, and NMP52 reporting sensitivity and specificity ranging from 82-94% and 80-93% respectively11, 15, 17, 20.
Four studies reported the diagnostic ability of proteins cytokeratin 8 and 18 using the UBC Rapid point of care Omega 100 reader. Cytokeratin are constituents of intermediate filaments of epithelial cells. This point of care test, requires three drops of urine and results from a photometric reader is available within 10 minutes. The sensitivity of the assay ranges from 30-87 % with carcinoma in situ (CIS) patients having the highest sensitivity and a specificity of between 63- 91% and a ROC of up to 0.750 suggesting a limited diagnostic performance. One study investigated the role of Ubiquitin 2 immunocytological staining reporting a sensitivity and specificity of 88% and 98% respectively although results for cytological based test were operator dependent.

A combination of urinary cytology, midkine (NEGF2) and gamma synuclein quantification using ELISA reported a ROC of 0.949 with a sensitivity and specificity of 91.8% and 97.5% respectively. The nonsulfated glycosaminoglycan hyaluronic acid (HA) quantified by ELISA reported a sensitivity and specificity of 88% and 82% increasing to 90% and 84% respectively when combined with hyaluronidase, a catalytic enzyme that degrades HA. Another 5-panel biomarker using gamma synuclein with Coronin-1A, Apolipoprotein A4, Semenogelin-2 and DJ-1/PARK7 compared ELISA to Western blot. Western blot achieved higher sensitivity (93.9% vs 79.2%) and a similar specificity (97% vs 100%) compared to ELISA in pTa/ pT1 cancers. However, western blot for protein quantification would not be practical in a large scale setting. Rosser et al. reported a RO of 0.948 using a multiplex ELISA system when combining three biomarkers: Interleukin 8 (IL-8), Matrix metallopeptidase 9 (MMP9) and vascular endothelial growth factor A (VEGFA). However, further studies incorporating the same three biomarkers and with the addition of between a further 4-7 markers have yielded an ROC of between 0.878-0.926 on validation studies.

Six studies utilised spectroscopy or chromatography to determine a metabolic signature or a molecular compound with a sensitivity and specificity of ≥ 80%. Several of these assays achieve sensitivity and specificities of ≥ 90% and while promising would require external validation.
Genomic biomarkers

Seven studies investigated the role of genomic biomarkers for the detection of bladder cancer. Four were based on analysis of mutations and included in Table A3. Telomerase reverse transcriptase (TERT) mutation represents the most common bladder cancer mutation present in > 70% of all bladder cancers. One study by Descotes and colleagues reported a sensitivity and specificity of 81% and 90% respectively for TERT although others have reported a lower sensitivity of 62%. TERT mutation was also associated with a > 5-fold increase relative risk of recurrence (p=0.0004).

FGFR3 achieved a sensitivity of 39% as a standalone test for bladder cancer. FGFR3 mutation is more common in low grade disease (p=0.02) and significantly associated with shorter time to recurrence (45% mutant vs 27% wild type, p=0.02). Other mutations such as TP53, PIK3CA and RAS have reported limited performance because of the low frequency of mutations and variability of genomic alterations between individual tumours. Sensitivity for TP53 of 12-13%, PIK3CA 13-14% and RAS 4.8% have been reported. The diagnostic performance of the combination of FGFR3 and TERT with PIK3CA, RAS and TP53 improved bladder cancer detection but only achieved a sensitivity of 73%. Of note, it has been demonstrated that following complete resection of tumour, 20.7% of patients will continue to test positive for FGFR3 and TERT mutation despite no cystoscopic detectable tumour in patients followed up for 3 years. In addition to targeted mutation analysis, the quantitative cell-free DNA analysis has been explored as a marker for the presence of bladder cancer as well as analysis of the integrity of cell-free DNA. To date studies are preliminary and report limited diagnostic performance with ROC of 0.725-0.834.

Epigenetic biomarkers

Twelve studies reported the diagnostic performance of microRNA (miRNA) and 8 studies investigated the role of DNA methylation as biomarkers for the detection of bladder cancer (Table 2 & A4). No studies investigated the role of histone modifications. Single target epigenetic biomarkers have a poor diagnostic performance overall and epigenetic biomarker panels with a sensitivity and
specificity of ≥ 80% are set out in Table 2. Of note, biomarker panels include between 2-150 targets to determine the presence of bladder cancer.

Of the miRNA panels, four have a sensitivity and specificity of ≥ 80% (Table 2) and employed miRNA arrays or next generation sequencing (NGS) to identify targets\textsuperscript{36, 38-40}. MiRNA was then quantified by real-time qPCR\textsuperscript{36-38, 40}. MiRNA-125b was used in two diagnostic panels although its sensitivity and specificity as a single biomarker varies between 59-85 and 76-96% respectively\textsuperscript{36, 74}. The combination of two miRNAs, miRNA-99a and miRNA-125b, had a sensitivity and specificity of 87% and 81% respectively\textsuperscript{36}. Using multivariable modeling Urquidi and colleagues determined the top 25 miRNA targets and determined the diagnostic ability of the top 10, 15, 20 and 25 targets using the LASSO approach to model the performance of each biomarker\textsuperscript{39}. Their results suggest that incorporating increasing number of biomarkers can increase both sensitivity and specificity with marginal gains with each increase.

Only three of the 8 DNA methylation studies reported sensitivity and specificity ≥ 80% (Table 2). All studies included ≥3 DNA methylation targets and all report a ROC of >0.9. Methylation status was determined by quantitative methylation specific PCR (qMS-PCR)\textsuperscript{42}, pyrosequencing\textsuperscript{41} and next generation sequencing\textsuperscript{43}. Su and colleagues interrogated three methylated targets and deduced that the combination of SOX1, IRAK3, L1-MET methylation had sensitivity and specificity of 80% and 97% respectively\textsuperscript{41}. The three-target methylation panel of POU4F2 + PCDH17 + GDF15 showed sensitivity and specificity of 91% and 88% respectively\textsuperscript{42}. Feber and colleagues derived a methylation signature of 150 loci incorporating a machine learning algorithm\textsuperscript{43}. The assay, UroMark, used a targeted bisulphite sequencing approach and was validated with two independent sets of urine samples comprising of bladder cancer and control samples reporting a sensitivity of 98%, specificity of 97% and ROC of 0.97\textsuperscript{43}.

Transcriptomic biomarkers

All studies used RT-PCR to determine expression of target genes (Table 3 & A5). Four studies report single target gene expression\textsuperscript{44-46, 53} and four studies combined
transcriptomic markers with urine cytology\textsuperscript{47, 48, 52, 53} to achieve a sensitivity and specificity of ≥ 80\% (Table 3). Of the four studies reporting a single biomarker, sensitivity ranges from 45-92\% and specificity of between 65-96\% and ROC of 0.741-0.966. Studies reporting combination biomarkers achieved a sensitivity of 36-97\%, specificity of 82-100\% and a ROC of 0.860-0.949.

S100A4, carbonic anhydrase IX (CAIX) and hepatoma upregulated protein RNA (HURP) and long non-coding RNA urothelial carcinoma associated-1 (lncRNA-UCA1) represent single biomarker targets which have sensitivity and specificity of ≥ 80\%\textsuperscript{44 45 46 53}. De Martino and colleagues quantified CAIX in paired tumour and urine and validated their results in an independent cohort comprising 155 urine samples reporting sensitivity, specificity and ROC of 81\%, 96\% and 0.883 respectively\textsuperscript{45}. Analysing six cytoplasmic calcium binding protein, S100A4 had the highest diagnostic accuracy with sensitivity of 90\%, specificity of 92\% and ROC of 0.978\textsuperscript{44}.

Eissa and colleagues used gold nanoparticle based RT-PCR and reported a sensitivity of 89\% and specificity of 94\% for the presence of Hepatoma upregulated protein RNA (HURP)\textsuperscript{46}. The technology performed better than conventional HURP RT-PCR, suggesting significant variation in results from different platforms\textsuperscript{47}. Another novel hybridization assay, nanoparticle RT-PCR of long non-coding RNA urothelial carcinoma associated-1 (lncRNA-UCA1) reported sensitivity and specificity of ≥ 90\% and ROC of 0.966\textsuperscript{53}. UCA1 has been implicated in bladder cancer progression through PI3K-AKT dependent pathways and the development of cisplatin resistance via Wnt signaling\textsuperscript{75, 76}. However, conventional RT-PCR of lncRNA-UCA1 has not reproduced these results\textsuperscript{77}.

Cytokeratin 20 (CK20) was used as part of two multiplex assays\textsuperscript{49, 52}. In contrast to CK8 and 18, CK20 is expressed on urothelium but not epithelial cells, and has a reported diagnostic sensitivity, specificity and ROC of 76-85\%, 86\% and 0.82-0.87 respectively\textsuperscript{49, 52}. CK20 overexpression in combination with p53 and Ki-67 have been shown by immunohistochemistry to suggest urothelial dysplasia\textsuperscript{78}. The combination of cytology with CK20 has a sensitivity and specificity of ≥ 90\% which has a higher diagnostic accuracy compared to other combinations such as Ki-67 with survivin, Ki-67 with CK20 and survivin with CK20\textsuperscript{49}. When CK20 is used in combination with
insulin like growth factor (IGF2), the sensitive and specificity increases to 90% and 84% respectively. The most promising transcriptomic panel that has been validated and tested in a prospective observational study is based on a combination of two genes IGF2 and Melanoma-associated antigen 3 (MAGE-A3). Both IGF2 and MAGE-A3 were selected from a panel of 12 genes and this two gene combination has a sensitivity of 81%, specificity of 91%, PPV of 87%, NPV of 88% and ROC of 0.944 in a prospective blinded validation study. The initial 12 gene expression targets were selected following screening using gene expression microarrays. IGF2 represents glycoprotein receptors on the cell membrane IGF2 promotes tumorigenesis via the PI3K-AKT pathway which is implicated in most bladder cancer. MAGE-A3 which has been shown to be expressed in 43% of bladder cancer and in various tumour types but not in healthy tissue with the exception of testis and placenta. Combination of different ‘omic’ biomarkers

Ten studies used a combination of difference ‘omic’ biomarkers with the aim to identify bladder cancer from exfoliated urinary bladder cells (Table 4 and Table A6). Six studies combined genomic with epigenetic biomarkers including one with microsatellite analysis. The other three studies used a transcriptomic and protein combination panel. One study utilised a protein (HYAL1), epigenetic (miR-210, miR-96) and transcriptomic (lncRNA-UCA1) combination. TERT and FGFR3 mutation were used in most combination markers incorporating genomic biomarkers. In a retrospective analysis of case control study of 74 bladder cancer and 80 controls presenting with haematuria, a combination of FGFR3, TERT and HRAS mutation in combination with twist-related protein (TWIST), OTX1 and ONECUT2 methylation, reported sensitivity of 97% and specificity of 83%. The authors modelled the PPV of 39% and NPV of 99.6% assuming a 10% prevalence of bladder cancer. This six gene panel of epigenetic and genomic targets, was subsequently validated in a prospective case control study with 97 bladder cancer and 103 controls presenting with haematuria with a sensitivity of 93% and ROC of 0.96. This assay builds on a
previously reported assay comprising of FGFR3 mutation in combination with OTX1, ONECUT2 and odd-skipped-related 1 (OSR1) methylation profile in a patient cohort of 95 cancer and 40 controls\textsuperscript{82}. This assay panel achieved a sensitivity of 79%, PPV of 92%, NPV of 76% and ROC of 0.864.

The other study by Dahmcke and colleagues was a prospective study with utilized a biomarker panel comprising of FGFR3 and TERT mutation with 6 methylated genes namely ONECUT2, Cyclin-A1 (CCNA1), BCL2, EOMES and vimentin (VIM)\textsuperscript{56}. This 8-biomarker combination had sensitivity of 97%, specificity of 76.9%, NPV of 99% and ROC of 0.963\textsuperscript{56}. Beukers and colleagues tested a three-panel biomarker comprising of FGFR3 and TERT mutation with OTX1 methylation and in pre-TURBT urine collection from 305 patients, achieving a sensitivity of 81-94% depending on tumour grade\textsuperscript{83}. However, in patients undergoing surveillance cystoscopy, the sensitivity and specificity of identifying tumour recurrence was much lower at 57-72% and 55-59% respectively\textsuperscript{83}.

A four-panel biomarker of FGFR3 mutation with Heparan sulfate glucosamine 3-O-sulfotransferase 2 (HS3ST2), SLIT2 or SEPTIN9 methylation was tested in a cohort of patients for the identification of NMIBC recurrence with surveillance cystoscopy\textsuperscript{86}. Roperch and colleagues incorporated clinical features such as age and smoking which improved the diagnostic accuracy of the assay from a sensitivity of 67-89% depending on tumour grade to 98% with an ROC of 0.96\textsuperscript{86}. However, when used in the surveillance setting, consistent with results from Beukers and colleagues, the sensitivity fell to 95% with an ROC of 0.82. Similarly, Zuiverloon and colleagues also observed that the diagnostic ability of urinary biomarkers to identify tumour recurrence during surveillance cystoscopy was poor\textsuperscript{84}.

The other three studies by Eissa et al. used combinations of protein and transcriptomics\textsuperscript{57-59}. Survivin involved in the EMT pathway was tested in combination with Matrix metalloproteinase (MMP) 2 & 9 and hyalurodinase. Survivin with MMP 2 & 9 had a sensitivity and specificity of 91% and 85% which increased to 96% and 85% when urinary cytology has been incorporated\textsuperscript{57}. Sensitivity and specificity of survivin with hyalurodinase was 95% and 90% respectively\textsuperscript{58}. The protein-epigenetic
combination of HYAL1, IncRNA-UCA1, \textit{miR-210} with \textit{miR-96} had a sensitivity of 100%, specificity of 89% and ROC of 0.981\textsuperscript{59}.

\textbf{Discussion}

This study highlights that single target assays have limited value regardless of ‘omic’ class. Performance is uniformly below that of multi-target biomarker panels. Only 4 single target urinary biomarkers achieved a sensitivity and specificity of ≥ 90% (Table 5). Across the studies none had a pre-planned statistical power calculation performed with only four non-case controlled prospective observational studies\textsuperscript{41, 50, 51, 56}. Independent validation cohorts were reported in six studies interrogating two biomarker panels. The first, a 10 protein based multiplex assay (IL8 + SERPINA1 + ANG + VEGF-A + CA9 + MMP 9 & 10 + APOE + PAI-1 + SDC1) and the second, a two panel gene expression assay (IGF2, MAGEA3)\textsuperscript{22-25, 50, 51}. Both assays reported a sensitivity and specificity of < 90% and ROC of <0.95. One panel comprising of 6 DNA methylation (\textit{SALL3 + ONECUT2 + CCNA1 + BCL2 + EOMES + VIM}) and two mutation (\textit{TERT & FGFR3}) was field tested in a prospective blinded patient cohort of haematuria patients reporting a sensitivity, specificity and ROC of 97%, 77% and
0.963 respectively but panel has not been validated in an independent patient cohort\textsuperscript{56}. A significant number of studies on urinary biomarkers had a poor diagnostic ability and require validation in a prospective clinical setting. Single and combination biomarkers with sensitive and specificity $\geq 80\%$ are shown in Table 5.

This study highlights that there is considerable interest in the use of urinary biomarkers to diagnose bladder cancer. This applies to both in the screening of the haematuria patient cohort as well as in patients with NMIBC who require surveillance cystoscopy. The requirement for cystoscopy represents a significant cost to health care services in diagnosing bladder cancer\textsuperscript{87}. Traditional imaging modalities with or without urine cytology does not have the necessary sensitivity to replace cystoscopy for the detection of bladder cancer \textsuperscript{88}. Cystoscopy requires a hospital visit and is an invasive procedure which is associated with a risk of urinary tract infection\textsuperscript{5}. A highly sensitive and specific non-invasive urinary assay will revolutionise both the haematuria and NMIBC surveillance pathway and is urgently needed.

In this study, we report that the diagnostic accuracy of urinary biomarkers varies considerably. In single target biomarkers had a sensitivity of 2-94\%, specificity of 46-100\%, PPV of 47-100\% and NPV of 21-94\%. Multi-target biomarkers achieved a sensitivity of 24-100\%, specificity of 48-100\%, PPV of 42-95\% and NPV of 32-100\%.

Such variation in diagnostic accuracy can be explained by combination of patient factors and assay factors. The diagnostic ability of urinary biomarkers was considerably better in identifying high grade tumours as well as CIS. This is constant with urinary cytology which has an overall 34\% sensitivity and 99\% specificity but the sensitivity increases to 63\% in CIS and high grade tumours\textsuperscript{89}. This is due to increase cell exfoliation in tumour cells and might in fact reflect why novel urinary biomarkers also detect high grade disease with a higher sensitivity and specificity. In fact advanced bladder cancer is often associated with a high mutational burden and hypermethylation\textsuperscript{90}.

Beside patient specific variables, reproducibility of biomarkers to allow highly accurate results is an issue. While efforts are made by the implementation of Good Laboratory Practice to uphold the quality of management controls to ensure
consistent and reliability of results, there are other sources of variation for the same biomarker. The variations in evaluating the same target protein, epigenetic change or gene expression makes it different to compare studies due to the lack of standardization of methodology. NGS performed in 5 different centers of the International Cancer Genome Consortium (IGGC) suggest that difference in variant calling and complete sequencing pipelines can result in a difference in identified mutation of ≥ 75%. Further, variation in genetic differences such as mutation, post transcription modifications, gene expression and epigenetic changes are complex and is difficult to elucidate. Additionally, the threshold used to define a positive result may differ between studies making comparison difficult.

A significant number of biomarkers reported did not have external validation in prospective field testing. For reasons described above, diagnostic accuracy of initial reports is often not reproducible. Where validation was performed, it was typically performed using selected patient cohort which is not representative of ‘real world practice’ of haematuria patients or NMIBC patients having surveillance cystoscopy. Majority of studies were based on retrospective patient cohorts comprising of selected bladder cancer and control patient groups. Hence, accurate PPV and NPV is not accurate or are based on assumptions as they are dependent on prevalence of disease in the patient cohort.

This study shows that the use of multi-target biomarkers is increasing and these biomarker panel have higher accuracy (Table 5). Traditionally, the number of biomarkers incorporated in an assay was limited by DNA yield from urinary cells. Female patients have a higher DNA yield compared to male patients. In addition, DNA extraction kit used and sampling time can also affect the DNA quality and yield. Particularly in methylation based assays which requires DNA bisulphite conversion, a loss of DNA yield of 70-90% is common. Fluorometer quantification of urinary DNA suggest that between 2 to 440 ng/ml of DNA can be retrieved from urinary cell pellet. In the studies reviewed, the limit on biomarker targets interrogated for protein, genomic, epigenetic, transcriptomic and combination biomarkers are 10, 5, 150, 12 and 8 respectively. The utility of NGS has allowed the development of highly multiplex assays, for genomic, epigenomic or transcriptomic...
biomarkers. The first to utilize this technology used multiplex biomarker panel of 150 loc\textsuperscript{43}.

The use of multi-target biomarkers is supported by seminal studies suggesting that there is significant intra-tumour heterogeneity within the same primary tumour\textsuperscript{95}. Hence, the diagnostic accuracy of biomarkers can be improved by a multitarget approach and it is unlikely that a single biomarker will be able to achieve a high diagnostic accuracy which meets the expectations of patients\textsuperscript{96}. While it is established that common mutations such as \textit{FGFR3} and \textit{TERT} are common in NMIBC, even in combination, a \textit{FGFR3-TERT} mutation assay will miss > 20\% of bladder cancers\textsuperscript{69}.

Currently, multi-panel biomarkers are often identified using next generation sequencing or arrays followed by a validation cohort of patients. However, incorporating more biomarkers may not improve diagnostic accuracy\textsuperscript{30, 50, 51}. The traditional methods such as defining a positive test using by a score and benchmarking it against an arbitrary threshold when evaluating multiple biomarkers is not ideal. Additionally, the choice of biomarkers to be incorporated is key. Using multiple biomarkers with a high sensitivity and specificity with significant overlap may risk poorer results. Hence, modern approaches incorporating complex bioinformatics and machine learning approaches using big data analysis represents a step change approach\textsuperscript{97}. Mathematical models such as random forest classifier or network models allows for the aggregation of higher sensitive and specific biomarkers with those of poorer accuracy that do not overlap resulting in a more robust test. In addition, considering KEGG pathways to determine truncal biological pathways implicated in bladder cancer carcinogenesis may allow for better biomarker selection which reflects functional biology\textsuperscript{98}. Further, aggregating different 'omic' biomarkers such as simultaneous analysis of DNA methylation, mutation, gene expression and copy number alterations has been hypothesized to improve biomarker accuracy\textsuperscript{99}. This approach has been utilised by two groups combining genomic with DNA methylation targets to achieve an ROC of 0.96\textsuperscript{55, 56}. Several studies also incorporated urinary cytology in addition to other biomarkers which resulted in improved biomarker performance\textsuperscript{20, 48, 52, 53}. Combining standard radiological images...
with genetic analysis has also proven to be an effective strategy in biomarker development\textsuperscript{100}.

The acceptable threshold of a urinary biomarker is dependent on its use as a companion test or a definitive test to replace cystoscopy. The NPV expected in a urinary assay used to replace cystoscopy in the hematuria setting is high given the devastating consequences in missing a bladder cancer particularly high-risk disease. In patient surveys, patients would only consider a urinary test with a diagnostic accuracy of $\geq 95\%$\textsuperscript{96}. However, when used as a companion test, currently available urinary biomarkers have been shown to increase the accuracy of cystoscopy which is operator dependent\textsuperscript{101}.

We acknowledge that there are limitations to our study. In our systematic review, we reviewed the published literature since 2013 hence reported markers with a high diagnostic accuracy published before 2013 will not be captured. However, given that no urinary biomarker still has the diagnostic ability to replace cystoscopy, we would expect that validation studies of promising biomarkers would continue to be reported. As with most studies, positive results are often reported, and negative results remain unpublished hence there might be more biomarkers investigated but they are likely to be of limited value.

The field of urinary biomarkers for the detection of bladder cancer is rapidly developing. However, no biomarkers reported today can replace cystoscopy. The lack of field testing, validation studies, use of different threshold to determine a positive test, tumour heterogeneity and complex interplay of different ‘omics’ represents challenges in in biomarker development and validation. However, NGS with the use of complex machine learning and mathematical modeling may represent a promising approach for biomarker discovery and promising biomarkers should be field tested to validate them.
Acknowledgements

We are grateful to the Medical Research Council (JDK, AF), Urology Foundation (WST, PK), Mason Medical Research Foundation (WST) & UCLH Biomedical Research Centre (JDK) for funding our work.


Table 1: Study characteristics and diagnostic accuracy of urinary protein biomarkers for the diagnosis of bladder cancer with sensitivity and specificity ≥ 80%.

| Title | Type of marker | Marker | Test/platform | Study design | Urine collection | Location | Country, % TCC | Low Grade (%) | Tumour arm | Control arm | Sensitivity | Specificity | PPV | NPV | ROC |
|-------|----------------|--------|---------------|--------------|-----------------|----------|---------------|---------------|-------------|-------------|-------------|-------------|-----------|-------|-----|-----|
| Li et al. 2016  | transport protein | ORM1 | ELISA | Case control | 20 ml midstream | China, 100% | 92 | 94 | 84 | 86 | 0.965 |
| Kato et al. 2014  | Hyaluronic acid | SDC1 | ELISA | Case control | 50 ml void | India, 100% | 56 | 59 | 80 | 81 | 0.861 |
| Chen et al. 2014  | DNA repair protein | APE1, RAD1 | ELISA | Case control | Not specified | Korea, 100% | 58 | 59 | 82 | 80 | 0.83 |
| Li et al. 2015  | Transcription regulator | ABI | ELISA | Case control | 50 ml midstream | China, Not specified | 42 | 54 | 52 | 51 | 0.527 |
| Lorenzo et al. 2013  | Nuclear protein | APOA1 | ELISA | Case control | Not specified | Italy, 100% | 68 | 68 | 81 | 82 | 0.84 |
| Liu et al. 2015  | HDL-related protein | APOC1 | ELISA | Case control | 50 ml midstream | China, Not specified | 102 | 106 | 85 | 84 | 0.848 |
| Ahn et al. 2014  | Nuclear matrix protein | NMP2A | ELISA | Case control | Not specified | Germany, 100% | 54 | 68 | 50 | 52 | 0.89 |
| Altahm et al. 2016  | Cell surface receptor | N-CAM | ELISA | Case control | Not specified | Egypt, Not specified | 56 | 68 | 54 | 56 | 0.91 |
| Song et al. 2015  | Heparan sulfate proteoglycan | HEP-1 | ELISA | Case control | Not specified | Japan, 100% | 100 | 100 | 80 | 80 | 0.94 |
| Jarsdulian et al. 2014  | Glycoconjugant | GM3 | ELISA | Case control | Not specified | Czech Republic, 100% | 27 | 70 | 98 | 98 | 0.9486 |
| Kumar et al. 2015  | Acute bleeding protein | Coagulation 1A | ELISA | Case control | 20 ml void | France, Germany, South Korea, Not specified | 98 | 95 | 80 | 72 | 0.91 |
| Hossner et al. 2014  | Chemokine (IL-8) | ELISA | Case control | 50 ml void | USA, Not specified | 100 | 100 | 81 | 81 | 0.904 |
| Goodson et al. 2016  | Chemokine (IL-8) | ELISA | Case control | 50 ml void | Japan, Not specified | 38 | 56 | 32 | 33 | 0.6925 |
| Srinivas et al. 2016  | Transcription factor | JUNB | ELISA | Case control | Not specified | Germany, Portugal, USA, Not specified | 17 | 100 | 85 | 85 | 0.912 |
| Hossner et al. 2016  | ELISA | Case control | >3 ml void | USA, Spain | 57 | 72 | 82 | 85 | 0.924 |
| Study et al. 2016 | Molecule signature | Reflection mode; Spectral range: 1500-1340, 1100-300, 900-800 | Infrared spectroscopy | Case control | 10 ml bladder wash | Cutoff | Not specified | Not specified | 40 | 21 | 82 | 81 | 90 | 81 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Nakai et al. 2015 | Porphyrin | Difference between ALA treated and ALA untreated samples at 935 nm | Fluorescence spectrophotometry | Case control | 15 ml void | Cutoff | Not specified | Not specified | 46 | 61 | 50 | 82 | 80 | 0.04 |
| Inoue et al. 2014 | Porphyrin | Fluorescence spectrophotometry | Case control | 15 ml void | Cutoff | Not specified | Not specified | 100 | 96 | 98 | 90 | 92 | 0.916 |
| Jin et al. 2014 | Metabolic signature | OPLAS-DA model: 12 peaks corresponding to succinate, pyruvate, oxoglutarate, carnitine, phosphoenolpyruvate, trimethyllysine, melatonin, taurine, carnitine, acetyl coA | Mass spectrometry | Case control | Morning void | Cutoff | Not specified | Not specified | 23 | 138 | 121 | 91 | 93 | 0.937 |
| Shen et al. 2015 | Metabolic signature | MixModel 1: GlyCysAlaLys, Inosinic acid, Trehalose, Nicotinuric acid, Asp Gly Trp, Ureidosuccinic acid | Mass spectrometry | Case control | Morning void | Cutoff | Not specified | Not specified | 23 | 21 | 91 | 81 | 93 | 0.934 |
| Aggio et al. 2016 | Metabolic signature | Principal component analysis | Mass spectrometry | Case control | 0.75 ml of morning void | Cutoff | Not specified | Not specified | 73 | 73 | 96 | 100 | 0.99 |
Table 2: Study characteristics and diagnostic accuracy of urinary epigenetic for the diagnosis of bladder cancer with sensitivity and specificity ≥80%.

<table>
<thead>
<tr>
<th>Title</th>
<th>Type of marker</th>
<th>Marker</th>
<th>Test platform</th>
<th>Study design</th>
<th>Urine collection</th>
<th>Country; % TCC</th>
<th>Low grade (%)</th>
<th>Tumour arm</th>
<th>Control arm</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
<th>ROC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zhang et al. 2014 (26)</td>
<td>miRNA</td>
<td>miR-99a + miR-125b</td>
<td>RT-qPCR</td>
<td>Case control</td>
<td>Urine supernatant</td>
<td>China, Not specified</td>
<td>30 50 21</td>
<td>87 85 92</td>
<td>71 0.975</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trissi et al. 2015 (27)</td>
<td>miRNA</td>
<td>MR-186, 188, 189, 190, 191, 192</td>
<td>RT-qPCR</td>
<td>Case control</td>
<td>Not specified</td>
<td>Egypt: 51-73</td>
<td>G3.2-73</td>
<td>94 60 87</td>
<td>86 80 0.927</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mengual et al. 2013 (28)</td>
<td>miRNA</td>
<td>18 miRNAs: miR-18a, miR-21a, miR-21b, miR-21c, miR-21d, miR-21e, miR-21f, miR-21g, miR-21h, miR-21i, miR-21j, miR-21k, miR-21l, miR-21m, miR-21n, miR-21o, miR-21p, miR-21q, miR-21r, miR-21s, miR-21t, miR-21u, miR-21v, miR-21w, miR-21x, miR-21y, miR-21z, miR-21aa, miR-21ab, miR-21ac, miR-21ad, miR-21ae, miR-21af, miR-21ag, miR-21ah, miR-21ai, miR-21aj, miR-21ak, miR-21al, miR-21am, miR-21an, miR-21ao, miR-21ap, miR-21aq, miR-21ar, miR-21as, miR-21at, miR-21au, miR-21av, miR-21aw, miR-21ax, miR-21ay, miR-21az, miR-21ba, miR-21bb, miR-21bc, miR-21bd, miR-21be, miR-21bf, miR-21bg, miR-21bh, miR-21bi, miR-21bj, miR-21bk, miR-21bl, miR-21bm, miR-21bn, miR-21bo, miR-21bp, miR-21bq, miR-21br, miR-21bs, miR-21bt, miR-21bu, miR-21bv, miR-21bw, miR-21bx, miR-21by, miR-21bz, miR-21ca, miR-21cb, miR-21cc, miR-21cd, miR-21ce, miR-21cf, miR-21cg, miR-21ch, miR-21ci, miR-21cj, miR-21ck, miR-21cl, miR-21cm, miR-21cn, miR-21co, miR-21cp, miR-21cq, miR-21cr, miR-21cs, miR-21ct, miR-21cu, miR-21cv, miR-21cw, miR-21cx, miR-21cy, miR-21cz, miR-21da, miR-21db, miR-21dc, miR-21dd, miR-21de, miR-21df, miR-21dg, miR-21dh, miR-21di, miR-21dj, miR-21dk, miR-21dl, miR-21dm, miR-21dn, miR-21do, miR-21dp, miR-21dq, miR-21dr, miR-21ds, miR-21dt, miR-21du, miR-21dv, miR-21dw, miR-21dx, miR-21dy, miR-21dz, miR-21ea, miR-21eb, miR-21ec, miR-21ed, miR-21ee, miR-21ef, miR-21eg, miR-21eh, miR-21ei, miR-21ej, miR-21ek, miR-21el, miR-21em, miR-21en, miR-21eo, miR-21ep, miR-21eq, miR-21er, miR-21es, miR-21et, miR-21eu, miR-21ev, miR-21ew, miR-21ex, miR-21ey, miR-21ez, miR-21fa, miR-21fb, miR-21fc, miR-21fd, miR-21fe, miR-21ff, miR-21fg, miR-21fh, miR-21fi, miR-21fj, miR-21fk, miR-21fl, miR-21fm, miR-21fn, miR-21fo, miR-21fp, miR-21fq, miR-21fr, miR-21fs, miR-21ft, miR-21fu, miR-21fv, miR-21fw, miR-21fx, miR-21fy, miR-21fz, miR-21ga, miR-21gb, miR-21gc, miR-21gd, miR-21ge, miR-21gf, miR-21gg, miR-21gh, miR-21gi, miR-21gj, miR-21gk, miR-21gl, miR-21gm, miR-21gn, miR-21go, miR-21gp, miR-21fq, miR-21fr, miR-21fs, miR-21ft, miR-21fu, miR-21fv, miR-21fw, miR-21fx, miR-21fy, miR-21fz, miR-21ga, miR-21gb, miR-21gc, miR-21gd, miR-21ge, miR-21gf, miR-21gg, miR-21gh, miR-21gi, miR-21gj, miR-21gk, miR-21gl, miR-21gm, miR-21gn, miR-21go, miR-21gp, miR-21f</td>
<td>40 miCtDNAs</td>
<td>30-50 ml etiologic</td>
<td>USA, Not specified</td>
<td>16 61 60</td>
<td>87 109</td>
<td>64 87</td>
<td>0.927</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gu et al. 2017 (29)</td>
<td>miRNA</td>
<td>7 cell-free miRNA: miR-7-3p, miR-122-3p, miR-29a-3p, miR-126-5p, miR-200a-3p, miR-37-5p, miR-451-5p, miR-421-5p</td>
<td>RT-qPCR</td>
<td>Case control</td>
<td>15 ml midstream urine, urine supernatant</td>
<td>China, Not specified</td>
<td>38 120 120</td>
<td>85 87</td>
<td>0.916</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Su et al. 2014 (30)</td>
<td>DNA methylation</td>
<td>DNA methylation</td>
<td>Bisulfite sequencing</td>
<td>Prospective cohort</td>
<td>50 ml void/bladder wash</td>
<td>USA, 100%</td>
<td>41</td>
<td>Not specified</td>
<td>89 57</td>
<td>97 26</td>
<td>0.91</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wang et al. 2016 (31)</td>
<td>DNA methylation</td>
<td>DNA methylation</td>
<td>qMS-PCR</td>
<td>Case control</td>
<td>Morning void</td>
<td>China; 100%</td>
<td>72 92</td>
<td>86 87 78</td>
<td>80 0.916</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>POU4F2</td>
<td>PCR</td>
<td>POU4F2</td>
<td>PCR</td>
<td>Case control</td>
<td>Morning void</td>
<td>USA; Not specified</td>
<td>91 59 60</td>
<td>94 80</td>
<td>0.919</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>POU4F2 + L1-MET</td>
<td>qPCR</td>
<td>L1-MET</td>
<td>qPCR</td>
<td>Case control</td>
<td>Morning void</td>
<td>USA; Not specified</td>
<td>91 59 60</td>
<td>94 80</td>
<td>0.919</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>POU4F2 + PCDH17 + GDF15</td>
<td>PCR</td>
<td>PCDH17 + GDF15 +</td>
<td>PCR</td>
<td>Case control</td>
<td>Morning void</td>
<td>USA; Not specified</td>
<td>91 59 60</td>
<td>94 80</td>
<td>0.919</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feber et al. 2017 (32)</td>
<td>DNA methylation</td>
<td>DNA methylation</td>
<td>RainDance</td>
<td>Case control</td>
<td>Voided urine</td>
<td>UK; Not specified</td>
<td>38 107 167</td>
<td>98 97</td>
<td>97 0.97</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

EDMES: Eomesoderm; GDF15: Growth/differentiation factor 15; IRAK3: Interleukin 1 Receptor Associated Kinase 3; L1-MET: Line 1 MET; NPV: negative predictive value; PPV: positive predictive value; PCDH17: Protocadherin-17; POU4F2: POU Class 4 Homeobox 2; TCC: transitional cell carcinoma; TCF21: Transcription factor 21
Table 3: Study characteristics and diagnostic accuracy of urinary transcriptomic biomarkers for the diagnosis of bladder cancer with sensitivity and specificity $\geq$ 80%.

<table>
<thead>
<tr>
<th>Study Design</th>
<th>Type of Markers</th>
<th>Tumour Arm</th>
<th>Control Arm</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>PPV (%)</th>
<th>NPV (%)</th>
<th>ROC AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case control</td>
<td>IGF2, MAGEA3, KLF9, CRH, SLC1A6</td>
<td>10-100 ml void</td>
<td>10-100 ml void</td>
<td>91%</td>
<td>87%</td>
<td>90%</td>
<td>85%</td>
<td>0.978</td>
</tr>
<tr>
<td>Case control</td>
<td>IGF2, MAGEA3, KLF9, CRH, SLC1A6</td>
<td>10-100 ml void</td>
<td>10-100 ml void</td>
<td>91%</td>
<td>87%</td>
<td>90%</td>
<td>85%</td>
<td>0.978</td>
</tr>
<tr>
<td>Prospective case-control</td>
<td>IGF2, MAGEA3, KLF9, CRH, SLC1A6</td>
<td>10-100 ml void</td>
<td>10-100 ml void</td>
<td>91%</td>
<td>87%</td>
<td>90%</td>
<td>85%</td>
<td>0.978</td>
</tr>
<tr>
<td>Prospective case-control</td>
<td>IGF2, MAGEA3, KLF9, CRH, SLC1A6</td>
<td>10-100 ml void</td>
<td>10-100 ml void</td>
<td>91%</td>
<td>87%</td>
<td>90%</td>
<td>85%</td>
<td>0.978</td>
</tr>
</tbody>
</table>

AHNAK2: AHNAK nucleoprotein 2; ANXA10: Annexin A10; CK20: Cytokeratin 20; CRH: Cortisol releasing hormone; CTSE: Cathepsin E; CTNNB1: Catenin beta 1; DEPTOR: DEPAG10; EBP: Early growth response protein; EBF1: Early B cell factor 1; EGF: Epidermal growth factor; FADD: FLICE-associated death domain protein; FABP5: Fatty acid binding protein 5; FGF19: Fibroblast growth factor 19; GADD45A: Growth arrest and DNA damage-inducible alpha; HUGP: Hepatitis-associated T cell growth factor; IGF2: Insulin like growth factor 2; KANK1: Knottin 1; KLF9: Krueppel like factor 9; KRT20: Keratin 20; MMP12: Matrix metalloproteinase 12; NPV: negative predictive value; PPV: positive predictive value; PPV/P1R4D: Protein phosphatase 1, regulatory (inhibitor) subunit 1D; SLC1A6: Solute carrier family 1 member 6; TCC: transitional cell carcinoma; TERT: Telomerase reverse transcriptase; TERT: Telomerase reverse transcriptase
<table>
<thead>
<tr>
<th>Title</th>
<th>Type of marker</th>
<th>Method</th>
<th>Type of test platform</th>
<th>Study design</th>
<th>Type of control</th>
<th>Study population</th>
<th>Sample size</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>TCC</th>
<th>NPV</th>
<th>PPV</th>
<th>Odds ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Van Kessel et al. 2016</td>
<td>Epigenetic + genomic</td>
<td>Methylation: TWIST1, CCNE1, BCL2, OTX1</td>
<td>Mutation analysis: FGFR3, TERT and HRAS</td>
<td>Case control</td>
<td>Not specified</td>
<td>Not specified</td>
<td>Not specified</td>
<td>20</td>
<td>74</td>
<td>80</td>
<td>97</td>
<td>83</td>
<td>23-39</td>
</tr>
<tr>
<td>Van Kessel et al. 2017</td>
<td>Epigenetic + genomic</td>
<td>Methylation: TWIST1, CCNE1, BCL2, OTX1</td>
<td>Mutation analysis: FGFR3, TERT and HRAS</td>
<td>Prospective case control</td>
<td>Not specified</td>
<td>Netherlands, Spain, Sweden</td>
<td>Not specified</td>
<td>26</td>
<td>97</td>
<td>103</td>
<td>93</td>
<td>86</td>
<td>0.56</td>
</tr>
<tr>
<td>Dahmcke et al. 2016</td>
<td>Epigenetic + genomic</td>
<td>Methylation: SALL3, CCNE1, CCNA1</td>
<td>miR96, LucRNA</td>
<td>Case control</td>
<td>Not specified</td>
<td>Denmark</td>
<td>100%</td>
<td>34</td>
<td>99</td>
<td>376</td>
<td>97</td>
<td>77</td>
<td>53</td>
</tr>
</tbody>
</table>

Table 4: Study characteristics and diagnostic accuracy of different combination ‘omic’ urinary biomarkers for the diagnosis of bladder cancer with sensitivity and specificity ≥80%.

BCL2: B-cell lymphoma 2; CCNA1: Cyclin A1; EOMES: Eomesodermin; FGFR3: fibroblast growth factor receptor 3; HYAL1: Hyaluronoglucosaminidase 1; lncRNA-UC1A: long non-coding RNA-prostate cancer associated 1; MMP2: matrix metalloproteinase-2; MMP9: matrix metalloproteinase-9; NPV: negative predictive value; ONECUT: One Cut Homeobox 2; OTX1: orthodenticle homeobox 1; PPV: positive predictive value; SALL3: spalt-like transcription factor 3; TCC: transitional cell carcinoma; TERT: Telomerase reverse transcriptase; TWIST1: Twist Family BHLH Transcription Factor 1; VIM: Vimentin
Table 5: Urinary biomarkers stratified according to ‘omic’ class and single vs multiple target biomarker with a sensitivity and specificity of ≥ 80%.

<table>
<thead>
<tr>
<th>Promising single biomarker</th>
<th>Protein</th>
<th>Genomic</th>
<th>Epigenetic</th>
<th>Transcriptomic</th>
<th>Promising biomarker combination</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>orosomucoid 1 (ORM1)*</td>
<td>TERT</td>
<td>POU Class 4 Homeobox 2*</td>
<td>S100A4</td>
<td>Amplified in breast cancer 1 + eukaryotic initiation factor 2 + Nuclear</td>
</tr>
<tr>
<td></td>
<td>Survivin</td>
<td></td>
<td></td>
<td>carbonic anhydrase IX</td>
<td></td>
</tr>
<tr>
<td></td>
<td>APE1/Ref-1</td>
<td></td>
<td></td>
<td>hepatoma upregulated protein RNA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Soluble FAS</td>
<td></td>
<td></td>
<td>Cytokeratin 20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HtrA1*</td>
<td></td>
<td></td>
<td>long non-coding RNA urothelial carcinoma associated-1*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Apo-A1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Calprotectin</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nuclear matrix protein 52</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ubiquitin 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hyaluronidase</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hyaluronic acid</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DJ-1/PARK7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Interleukin-8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>uroporphyrin I</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>coproporphyrin</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AIB1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Matrix Protein 22 | Apolipoprotein A1 + cytology  
| Corin-1A + Apolipoprotein A4  
| Hyaluronic acid + hyaluronidase  
| Coronin-1A + Apolipoprotein A4 + Semenogelin-2 + synuclein-g + PARK7/DJ-1*  
| Interleukin 8+ Matrix metallopeptidase 9 + Vascular endothelial growth factor A  
| Interleukin 8 + SERPINA1 + ANG + Vascular endothelial growth factor A  
| + CA9 + Matrix metallopeptidase 9 & 10 + Apolipoprotein E  
| Plasminogen activator inhibitor-1+ Syndecan†  
| Spectral range- 1500-1340, 1100-900, 900-800  
| Metabolic signature- succinate, pyruvate, oxoglutarate, carnitine, phosphoenolpyruvate, trimethyllysine, melatonin, isavalsrylcarnitine, glytarylcarnitine, octenoylcarnitine, decanoylcarnitine, acetyl-coA*  
| Metabolic signature- GlyCysAlaLys, Inosinic acid, Trehalose, Nicotinuric acid, Asp Asp Gly Trp, Ureidosuccinic acid  
| Principal component analysis*  
|  
| Epigenetic | mRNA-99a +mRNA-125b  
| MiR-96+ cytology  
| miR-187 + miR-18a + miR-25 + miR-142-3p + miR-140-5p + miR-204  
| 10 and 25 panel miR  
| Cell free: miR-7-5p + miR-22-3p + miR-29a-3p + miR-126-5p + miR-200a-3p + miR-375 + miR-423-5p  
| methylation: SOX1 + Interleukin 1 Receptor Associated Kinase 3 + Line 1 MET  
| methylation: POU Class 4 Homeobox 2 + Protocadherin-17*  
| methylation: 150 CpG sites*  
|  
| Transcriptomic | hepatoma upregulated protein + cytology*  
| X-linked inhibitor of apoptosis protein + cytology*  

---

*Indicates significant alterations in expression levels.
| Multi ‘omic’ biomolecule | • Cytokeratin 20 + cytology*  
• Survivin + cytology*  
• Ki67 + Cytokeratin 20  
• Insulin like growth factor 2, Melanoma-associated antigen 3  
• Cytokeratin 20 + Insulin like growth factor 2  
• long non-coding RNA urothelial carcinoma associated-1 + cytology*  

| Methylation: Twist Family BHLH Transcription Factor 1, One Cut Homeobox 2 + orthodenticle homeobox 1. Mutation: Fibroblast growth factor receptor 3, Telomerase reverse transcriptase and HRAS  
| Matrix metalloproteinase 2 & 9 (protein) + survivin (mRNA) + cytology*  
| Survivin (protein) + hyaluronidase (mRNA) + cytology*  
| HYAL1 (protein) + miR-210 + miR96+ long non-coding RNA-urothelial cancer associated 1 (mRNA) + cytology*  

*≥90% sensitivity and specificity  
\[1\] independent cohort validation studies
Figure 1: Flow chart of studies identified, excluded and included.

- Total number of studies identified through database searching, n=646
- Additional records identified from other sources, n=11

- Abstracts screened, n=657
  - Total studies excluded, n=377
    - Non-English, n=30
    - Literature reviews, n=111
    - Editorials/letters, n=16
    - Not related, n=221

- Full text screened, n=279
  - Total studies excluded, n=164
    - Commercial test, n=69
    - <20 patients in each arm, n=38
    - Prognosis, n=16
    - Single arm, n=26
    - No sensitivity/specificity/AUC data, n=15

- Included articles, n=115