Urine Markers as Possible Tools for Prostate Cancer Screening: Review of Performance Characteristics and Practicality

Heiko Müller and Hermann Brenner

Background: In recent years, an increasing number of urine-based tests have been proposed as potential screening tests for prostate cancer. The goal of this review was to summarize the current status of evidence regarding performance characteristics of the proposed tests and their practicality under screening conditions.

Method: Relevant articles published up to and including May 2005 were identified in the PubMed database. At least 10 cases and 10 controls had to be analyzed for a study to be included in the review. Data concerning the study population, performance characteristics, and the collection and processing of urine samples were extracted from the reviewed articles.

Results: In all, 34 retrospective studies evaluating 21 different markers complied with the inclusion criteria. Most of the studies were rather small and included heterogeneous clinical study populations. Promising results were reported for a few markers in single studies, but they have often not been replicated in subsequent larger studies. Some of the more promising results were obtained with 24-h urines or with specimen-handling procedures that might be difficult to perform under screening conditions.

Conclusions: Larger studies with a prospective design are required to confirm promising findings regarding performance characteristics of some novel markers recently reported in mostly small studies. Future studies should also pay particular attention to the practicality of the markers under screening conditions.

Prostate cancer is the most common cancer in men globally, with an estimated prevalence of more than 2 million cases. In the year 2002, a total of 221,000 men died of prostate cancer worldwide (1). Prostate cancer develops more slowly than most other cancers, which favors the development of a successful screening program to reduce mortality and morbidity. The predominant current screening methods include serum prostate-specific antigen (PSA) measurements and digital rectal examination (DRE) with a subsequent transrectal ultrasound-guided biopsy of the prostate if either of these 2 methods shows a positive result. Because the performance of serum PSA is superior to that of DRE (2), testing for PSA has become widespread in an increasing number of countries. However, the usefulness of such testing has been the subject of ongoing controversy (3). In particular, a substantial proportion of men with prostate cancer have PSA values less than the commonly used cutpoint of 4 μg/L (4), and many unnecessary biopsies are performed because of high PSA values (5). It thus is highly desirable to develop reliable, easily performed, and widely accepted screening tests with better performance characteristics. Noninvasive urine-based tests might be particularly interesting candidates for that purpose.

The purpose of urine-based screening tests for prostate cancer is to find cancer cells from which markers can be extracted or to find released proteins or nucleic acids that are modified compared with the forms in healthy men. Three different groups of markers can thus be considered for the detection of prostate cancer: DNA-, RNA-, and protein-based markers. The main challenge is to find a marker that has good performance characteristics and at the same time allows easy urine collection and processing.

The primary aim of this review was to summarize the current evidence regarding performance characteristics of tests proposed for urine-based prostate cancer detection.
Another important aspect was to examine the practicality of these markers under screening conditions.

Materials and Methods
Pertinent studies were searched in the PubMed database, including publications up to and including May 2005. The following terms were used in different combinations to identify relevant reports: “prostate cancer”, “urine”, “urin*”, “test”, “marker”, “biomarker”, “diagnosis”, “screening”, “DNA”, “RNA”, and “protein”. The bibliographies of the identified publications were also considered.

Only human studies published in English were considered for the review. We used sensitivity and specificity to describe the performance characteristics of the tests; therefore, we included only studies that contained both prostate cancer cases and controls. To ensure a minimum precision of estimates of sensitivity and specificity, we included only studies with at least 10 cases and 10 controls. We calculated the 95% confidence intervals of those parameters based on the exact binomial distribution.

Among the studies complying with the inclusion criteria, 3 major different groups were distinguished according to the type of marker used: DNA-, RNA-, or protein-based markers. Among the DNA-based markers, a further distinction was made between genetic and epigenetic markers. Among protein-based markers, a further distinction was made between qualitative and quantitative markers. For quantitative markers, the cutoff point used in each study is given in one of the tables in this review. In a few articles, the cutoff point was not explicitly mentioned but could be derived from a figure. Information concerning the study population, performance characteristics, and collection and processing of urine markers was extracted for the review.

The numbers of cases and controls as well as mean or median ages were extracted to describe the study population. If possible, Gleason scores and tumor stage or grade were included. If not otherwise mentioned, cases were mostly confirmed by biopsy, which is the current gold standard for confirmation of prostate cancer. In addition, whenever possible, the criteria used for selection of controls were recorded. It was not clear in all reports whether the controls were confirmed to be free of prostate cancer by biopsy. If this information was reported, it is given one of the tables in this review. In some studies, women were included as controls. Whenever possible, specificity was calculated after exclusion of women from the controls. If this was not possible, the original specificity including women was used for the review and the number of women is given in the table.

Information concerning collection and processing of urine samples was extracted to assess the practicality and suitability of the marker for possible use as a screening tool. For example, a few studies in which urine was collected by urethral washing or milking or by catheter were included in the review, and this is listed in the tables.

Finally, we looked at potential sources of biases, e.g., whether the authors of the report explicitly indicated that those who performed the analyses were blinded to the diagnosis, whether cases and controls were comparable with respect to age and other factors that might affect test performance, and whether urine sampling procedures were similar for cases and controls.

Results
Among 60 initially identified studies of potential interest, 34 studies complied with all inclusion criteria. Overall, 21 different markers were analyzed in these studies. Studies were excluded for the following reasons: <10 cases and/or 10 controls (12 studies); sensitivity and/or specificity not mentioned and not calculable from the given information (13 studies); and article not in English (1 study). In 12 studies, descriptive information was available concerning Gleason score, stage, or grade of tumor. In 3 studies, the sensitivity was calculated according to tumor stage or Gleason score. The association between the tumor marker and Gleason score, tumor stage, or grade was mentioned in 6 studies.

Studies evaluating genetic alterations are listed in Table 1. Two studies investigated loss of heterozygosity (LOH) at defined locations as a tumor marker (6, 7). LOH might be the most common deletion event in prostate cancer (8). Cussenot et al. (6) assessed 4 locations and obtained a sensitivity of 73% and a specificity of 67% for LOH at one or more of the locations. Including 2 additional locations, Thuret et al. (7) obtained a sensitivity of 87% and a specificity of 44%. Urine was collected after prostatic massage in both studies.

A genetic biomarker of cellular oxidative stress that might be related to cancer is 8-hydroxydeoxyguanosine. Using a cutoff of 100 μg/g of creatinine, Chiou et al. (9) calculated a sensitivity of 31% and a specificity of 100% for the detection of prostate cancer by use of this marker.

Studies assessing epigenetic alterations are summarized in Table 2. All 5 studies (10–14) evaluated the performance characteristics of promoter hypermethyl- ation of the glutathione S-transferase P1 gene as a tumor marker. This DNA alteration appears in >90% of prostatic carcinoma tissues. Sensitivity was between 19% and 76%, and specificity ranged from 56% to 100%. Sensitivity was lowest (19%–30%) in the only study in which urine was collected without previous prostatic massage or previous biopsy. Where assessed, no significant association with either Gleason score (12, 13) or tumor stage (10, 11) was found.

Studies evaluating RNA-based urine markers are summarized in Table 3. The highest number of specimens analyzed for a particular marker was for DD3PCA3 after prostate massage, which was analyzed in 3 different studies (15–17). A noncoding messenger RNA is expressed by the DD3PCA3 gene in epithelial prostate cells
and is overexpressed in prostate cancer tissue samples compared with nonmalignant tissue (16). The sensitivity in the 3 studies ranged from 66% to 82% and the specificity from 76% to 89%. Unfortunately, comparison of age between cases and controls was not possible with the data reported in the 3 studies. Furthermore, only samples expressing enough PSA were included in the studies; therefore, probably only persons at higher risk were evaluated in these studies.

Cells may escape senescence and proliferate if telomerase is activated (18). Expression of human telomerase reverse transcriptase (hTERT) is critical for telomerase activity. In their study, Crocitto et al. (14) measured hTERT RNA expression by reverse transcription-PCR and obtained a sensitivity and specificity of 36% and 66%, respectively (Table 3). Using different approaches, Meid et al. (18) and Vicentini et al. (19) measured telomerase activity directly with the telomeric repeat amplification protocol assay and obtained sensitivities of 58% and 90% and specificities of 100% and 87%, respectively (Table 4). In all 3 studies, prostatic massage was performed. Furthermore, Meid et al. (18) found a significant association between Gleason score and telomerase activity.

The performance characteristics of survivin, an inhibitor of apoptosis (20), as a marker for prostate cancer was evaluated in 2 studies (20, 21). Whereas Wang et al. (20) measured mRNA expression (RNA-based approach; Table 3), Smith et al. (21) used a polyclonal antibody to detect survivin (protein-based approach; Table 4). Sensitivity was 0% (whereas it was 100% and 80% for bladder cancer) in both studies, and specificity reached 100% (20) and 91% (21), respectively.

An overview of studies evaluating protein-based quantitative markers, including the studies of telomerase (18, 19) and survivin (21) discussed above, is given in Table 4. Prostatic inhibin-like peptide is involved in the suppression of follicle-stimulating hormone (22). In each of 2 studies, both based on collection of 24-h urines and using a slightly different cutpoint, Teni and coworkers (22, 23) estimated a sensitivity >80% and specificity of 100%; however, these impressive results, obtained in 1988 and 1989, have not subsequently been reproduced.

Stoeber et al. (24) assessed minichromosome maintenance 5 (MCM-5) protein as a potential marker for prostate cancer. Minichromosome maintenance proteins are involved in the initiation of DNA replication and thus play a critical regulatory role (24). The high estimate of sensitivity (92%) in this blinded study was based on rather small numbers of patients, but a relatively precise estimate of 82% based on more than 200 controls was obtained for specificity. Two other markers, bladder tumor fibronectin (25) and basic human arginine amidase (26), have also been evaluated, each in a single study. It was not clear in the report regarding bladder tumor fibronectin (25) whether there was a difference in urine sampling between cases and controls because voided and catheterized samples were collected. Otherwise, there
Table 2. DNA-based markers (epigenetic alterations).

<table>
<thead>
<tr>
<th>Authors, year (reference)</th>
<th>Marker</th>
<th>Study population</th>
<th>Cases</th>
<th>Controls</th>
<th>Sensitivity, n (%)a</th>
<th>Specificity, n (%)a</th>
<th>Collection of urine samples</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goessl et al., 2001 (10)</td>
<td>Promoter hypermethylation of GSTP1b</td>
<td>40 with Pca (stage T1–T2, n = 22; stage ≥T3, n = 18); median age, 65 years</td>
<td>45 with BPH (31 by biopsy); median age, 62 years</td>
<td>29 of 40 [73 (56–85%); for stage T1–T2, 68%; for stage ≥T3, 78%]</td>
<td>44 of 45 [98 (88–100%)]</td>
<td>After prostatic massage; centrifugation</td>
<td>No significant association between tumor stage and detection of neoplastic DNA</td>
<td></td>
</tr>
<tr>
<td>Goessl et al., 2001 (11)</td>
<td>Promoter hypermethylation of GSTP1</td>
<td>11 with Pca [mean age, 68 years; group 1 (see comments)]; 29 with Pca [mean age, 68 years; group 2 (see comments)]</td>
<td>10 with BPH, [histology; mean age, 64 years (group 1)]; 40 with BPH [mean age, 62 years (group 2)]</td>
<td>Group 1, 4 of 11 [36 (11–69%)]; group 2, 22 of 29 [76 (56–90%)]</td>
<td>Group 1, 10 of 10 [100 (69–100%)]; group 2, 39 of 40 [97 (87–100%)]</td>
<td>After prostatic massage</td>
<td>Group 1, DNA extraction by QiAamp Blood and Tissue Kit; only 21 of 31 patients with enough DNA; group 2, DNA extraction by QiAamp RNA Viral Kit; no sign of association between GSTP1 and tumor stage</td>
<td></td>
</tr>
<tr>
<td>Jeronimo et al., 2002 (12)</td>
<td>Promoter hypermethylation of GSTP1</td>
<td>69 with localized Pca; median age, 63 years</td>
<td>31 with BPH (resection); median age, 64 years</td>
<td>21 of 69 [30 (20–43%); by conventional MSP; 9 of 69 [19 (6–23%); by RTQ MSP]</td>
<td>30 of 31 [97 (83–100%); by conventional MSP; 30 of 31 [97 (83–100%); by RTQ MSP]</td>
<td>To laboratory on ice, centrifuged, stored at −80 °C</td>
<td>No sign of association between GSTP1 methylation status and Gleason score</td>
<td></td>
</tr>
<tr>
<td>Gonzalgo et al., 2003 (13)</td>
<td>Promoter hypermethylation of GSTP1</td>
<td>12 with Pca; mean age, 67.8 years</td>
<td>18 with no cancer on biopsy; mean age, 62.6 years</td>
<td>7 of 12 [58 (28–85%)]</td>
<td>11 of 18 [61 (36–83%)]</td>
<td>After biopsy; stored at 4 °C and centrifuged</td>
<td>Blinded analysis; no sign of an association between methylation of GSTP1 and Gleason score; 9 specimens did not have sufficient DNA for PCR</td>
<td></td>
</tr>
<tr>
<td>Crocitto et al., 2004 (14)</td>
<td>Promoter hypermethylation of GSTP1</td>
<td>24 with Pca; age (see comments)</td>
<td>34 with negative biopsy; age (see comments)</td>
<td>11 of 24 [46 (26–67%)]</td>
<td>19 of 34 [56 (38–73%)]</td>
<td>Urethral milking after prostate massage; buffered with phosphate, pelleted, and stored at −80 °C</td>
<td>Age range over all groups was 48–83 years</td>
<td></td>
</tr>
</tbody>
</table>

a Percentage values in parentheses are the 95% confidence interval.
b GSTP1, glutathione S-transferase P1 gene; Pca, prostate cancer; BPH, benign prostatic hyperplasia; MSP, methylation-specific PCR; RTQ, real-time quantitative.
Table 3. RNA-based markers.

<table>
<thead>
<tr>
<th>Authors, year (reference)</th>
<th>Marker (cutoff value)</th>
<th>Study population</th>
<th>Sensitivity, n (%)a</th>
<th>Specificity, n (%)a</th>
<th>Collection of urine samples</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hessels et al., 2003 (15)</td>
<td>Ratio of DD3PCA3 to urinary PSA (200 copies of DD3 per 1000 copies of PSA)</td>
<td>24 with Pca; age not mentioned</td>
<td>16 of 24 [67 (45–84)]%</td>
<td>70 of 84 [83 (74–91)]%</td>
<td>After prostatic massage; cooled on ice; centrifugation; phosphate buffered, stored at −70 °C</td>
<td>Blinded analysis; PSA &gt; 3</td>
</tr>
<tr>
<td>Tinzi et al., 2004 (16)</td>
<td>DD3PCA3 calculated probability of cancer (&gt;0.5)</td>
<td>62 with Pca; age (see comment); Gleason score &lt;7 (n = 37), 7 (n = 14), &gt;7 (n = 11)</td>
<td>51 of 62 [82 (70–91)]%</td>
<td>72 of 95 [76 (66–84)]%</td>
<td>First urine after rectal palpation; buffered with phosphate, stored at 4 °C, and processed within 48 h; centrifugation + buffer, then stored at −20 °C</td>
<td>Only age range in all groups (41–89 years) mentioned; 43 samples with insufficient prostate cells for analysis</td>
</tr>
<tr>
<td>Fradet et al., 2004 (17)</td>
<td>DD3PCA3 by classification tree</td>
<td>152 with Pca; age (see comment); Gleason score 6–7 (72%)</td>
<td>100 of 152 [66 (58–73)]%</td>
<td>258 of 291 [89 (84–92)]%</td>
<td>First voided urine after palpation; buffered with phosphate and shipped at 2–8 °C to laboratory</td>
<td>Only median age (64 years) of all groups mentioned; 74 samples had PSA too low for analysis; multicenter study</td>
</tr>
<tr>
<td>Crocitto et al., 2004 (14)</td>
<td>hTERT expression</td>
<td>14 with Pca; age (see comment); median Gleason score, 6</td>
<td>5 of 14 [36 (13–65)]%</td>
<td>23 of 35 [66 (48–81)]%</td>
<td>Urethral milking after prostate massage; most samples buffered with phosphate and stored at −80 °C</td>
<td>Age range over all groups was 48–83 years</td>
</tr>
<tr>
<td>Wang et al., 2004 (20)</td>
<td>Survivin mRNA expression</td>
<td>10 with Pca; age not mentioned</td>
<td>0 of 10 [0 (0–29)]%</td>
<td>20 of 20 [100 (83–100)]%</td>
<td>200 mL of urine; centrifuged; RNA extracted and stored at −80 °C</td>
<td>80% sensitivity for bladder cancer (n = 20)</td>
</tr>
</tbody>
</table>

a Percentage values in parentheses are the 95% confidence interval.

b DD3, differential display code 3; Pca, prostate cancer; PIN, prostatic intraepithelial neoplasia; ASAP, atypical small acinar proliferation; BPH, benign prostatic hyperplasia.
Table 4. Protein-based quantitative urine markers.

<table>
<thead>
<tr>
<th>Authors, year (reference)</th>
<th>Marker (cutoff value)</th>
<th>Study population</th>
<th>Cases</th>
<th>Controls</th>
<th>Sensitivity, n (%)</th>
<th>Specificity, n (%)</th>
<th>Collection of urine samples</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meid et al., 2001 (18)</td>
<td>Telomerase activity by TRAP(^a) assay</td>
<td>16 with Pca (radical prostatectomy + fresh voided urine); 8 with Pca (biopsy + UW); age see comment; median Gleason score, 7</td>
<td>12 with BPH (biopsy + UW); age see comment</td>
<td>Fresh urine + UW, 14 of 24 [58 (37–78)%]; UW only, 4 of 8 (50%); fresh urine only, 10 of 16 (63%); poor differentiation, 40%; good or moderate differentiation, 100%</td>
<td>Fresh voided urine or UW after prostatic massage; buffered, centrifuged, stored at (-80 °C)</td>
<td>Overall age was 50–86 years in all groups; Gleason score significantly associated with telomerase activity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vicentini et al., 2004 (19)</td>
<td>Telomerase activity by TRAP assay</td>
<td>30 with Pca; age (see comment); Gleason score (\leq5) ((n = 13), 6) ((n = 6), 7) ((n = 11))</td>
<td>30 with BPH; age (see comment); histologically confirmed</td>
<td>27 of 30 ([90 (73–98)%]); Gleason score (\leq5) ((84%), 6) ((100%), 7) ((91%))</td>
<td>26 of 30 ([87 (69–96)%])</td>
<td>First urine after prostatic massage; centrifuged and stored at (-80 °C) (pelleted cells)</td>
<td>Blinded analysis; all participants older than 50 years</td>
<td></td>
</tr>
<tr>
<td>Smith et al., 2001 (21)</td>
<td>Survivin ((0.001 mg/L))</td>
<td>19 with Pca; age not mentioned</td>
<td>16 healthy volunteers (mean age, 47.7 years); 29 with nonneoplastic urinary tract diseases (mean age, 60 years)</td>
<td>0 of 19 ([0 (0–18)%])</td>
<td>41 of 45 ([91 (79–98)%])</td>
<td>Random clean-catch or straight catheter urine, stored at (-80 °C) and then centrifuged; supernatant filtered</td>
<td>100% sensitivity for bladder cancer ((n = 31))</td>
<td></td>
</tr>
<tr>
<td>Teni et al., 1989 (22)</td>
<td>Prostatic inhibin-like peptide ((58.2 \mu g/24-h urine))</td>
<td>21 with Pca; age range, 50–90 years; grades III/IV</td>
<td>20 healthy controls; age range, 45–85 years</td>
<td>17 of 21 ([81 (58–95)%])</td>
<td>20 of 20 ([100 (83–100)%])</td>
<td>24-h urine kept at 4 °C during collection; aliquots stored at (-20 °C)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teni et al., 1988 (23)</td>
<td>Prostatic inhibin-like peptide ((60 \mu g/24-h urine))</td>
<td>25 with Pca; age range, 50–85 years; grades III/IV</td>
<td>24 healthy controls (age range, 46–80 years) 17 with BPH (age range, 50–85 years)</td>
<td>21 of 25 ([84 (64–95)%])</td>
<td>41 of 41 ([100 (91–100)%])</td>
<td>24-h urine kept at 4 °C during collection and then stored at (-20 °C)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stoeber et al., 2002 (24)</td>
<td>MCM-5 protein ((1500 HeLa cells/well))</td>
<td>12 with Pca; ages (see comment)</td>
<td>70 with BPH; 131 with no BPH; ages (see comment)</td>
<td>11 of 12 ([92 (62–100)%])</td>
<td>165 of 201 ([82 (76–87)%])</td>
<td>Urine centrifuged, buffered, and stored in liquid nitrogen</td>
<td>Blinded analysis; age range, 62–78 years for all participants</td>
<td></td>
</tr>
<tr>
<td>Sanchez-Carbayo et al., 2000 (25)</td>
<td>Bladder tumor fibronectin ((52.8 \mu g/L))</td>
<td>21 with Pca; mean age, 70.2 years</td>
<td>18 healthy males (mean age, 63.0 years); 13 with BPH (mean age, 70.2 years)</td>
<td>9 of 21 ([43 (22–66)%])</td>
<td>24 of 31 ([77 (59–90)%])</td>
<td>Voided or catheterized urine, centrifuged and stored at (-20 °C)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Matsuda et al., 1996 (26)</td>
<td>BHUAEl activity 15 nmol/(min x day) (^c)</td>
<td>15 with Pca; ages (&gt;59) years</td>
<td>12 healthy volunteers (ages (&gt;55) years); 6 with BPH (ages (&gt;59) years)</td>
<td>7 of 15 ([47 (21–73)%])</td>
<td>18 of 18 ([100 (81–100)%])</td>
<td>Urine stored frozen until used, dialyzed, buffered, and centrifuged</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Authors, year (reference)</td>
<td>Marker (cutoff value)</td>
<td>Study population</td>
<td>Sensitivity, n (%)</td>
<td>Specificity, n (%)</td>
<td>Collection of urine samples</td>
<td>Comments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
<td>-----------------------</td>
<td>------------------</td>
<td>-------------------</td>
<td>-------------------</td>
<td>-----------------------------</td>
<td>----------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joseph et al., 1995 (27)</td>
<td>Ratio of scatter factor to creatinine ratio (15 pg/mg creatinine)</td>
<td>23 with Pca; ages not mentioned</td>
<td>7 of 23 (30 [13–53%])</td>
<td>17 of 18 [94 (73–100%)]</td>
<td>Stored at −20 °C, centrifuged; supernatant subjected to ultrafiltration</td>
<td>Stored at −20 °C, centrifuged; supernatant used with acetic acid and stored at −20 °C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rosen et al., 1997 (28)</td>
<td>Ratio of scatter factor to creatinine (20 pg/mg creatinine)</td>
<td>49 with Pca; ages not mentioned</td>
<td>15 of 49 (31 [18–45%])</td>
<td>36 of 38 [94 (82–99%)]</td>
<td>Midstream urine stored at −20 °C, centrifuged, and subjected to ultrafiltration</td>
<td>Midstream urine stored at −20 °C, centrifuged, and subjected to ultrafiltration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fernandez et al., 1986 (29)</td>
<td>Transferrin (7.5 mg/L); see comment</td>
<td>22 with Pca; age range, 46–87 years; advanced Pca</td>
<td>8 of 22 (36 [17–59%])</td>
<td>&gt;95% (see comment)</td>
<td>Urine mixed with buffered transferrin antiserum</td>
<td>Cutoff value chosen higher than upper limit of transferrin values from persons with BPH and much higher than upper limit of 95% confidence interval of persons with nonmalignant disorders (only 95% confidence interval was provided)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dieijen-Visser et al., 1988 (30)</td>
<td>Ratio of transferrin to creatinine (0.1 g/mol creatinine)</td>
<td>34 with Pca; ages not mentioned; stage T1 (n = 6), T2 (n = 7), T3 (n = 19), T4 (n = 2)</td>
<td>75% (see comment)</td>
<td>30% (see comment)</td>
<td>Urine samples stored at −70 °C, centrifuged, and adjusted to pH 7</td>
<td>Estimated from the ROC curve</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abdul and Hoosein, 1995 (31)</td>
<td>(IgG + IgA + IgM) concentrations (~63 000 cpm bound)</td>
<td>30 with Pca; age not mentioned; stage C (n = 8), D1 (n = 8), D2 (n = 14)</td>
<td>10 of 30 (13 [17–53%]; stage C, 13%; D1, 38%; D2, 43%)</td>
<td>15 of 15 [100 (78–100%)]</td>
<td>Cases, first-morning urine; controls, collection not mentioned; centrifuged, supernatant used with acetic acid and stored at −20 °C</td>
<td>Cases, first-morning urine; controls, collection not mentioned; centrifuged, supernatant used with acetic acid and stored at −20 °C</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4. Continued

<table>
<thead>
<tr>
<th>Authors, year (reference)</th>
<th>Marker (cutoff value)</th>
<th>Study population</th>
<th>Controls</th>
<th>Sensitivity, n (%)&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Specificity, n (%)&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Collection of urine samples</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lwaleed et al., 2000 (32)</td>
<td>Urinary tissue factor (12.3 μg/L)</td>
<td>26 with Pca; median age, 74 years</td>
<td>57 healthy persons (27 female, 30 male; median age, 33 years); 30 patients with renal stones (16 female, 14 male; median age, 44 years); 67 patients with BPH (median age, 72 years)</td>
<td>17 of 26 [65 (44–83)%]</td>
<td>115 of 154 [75 (67–81)%]</td>
<td>Centrifugation and solubilization</td>
<td>Significant increase in urinary tissue factor concentrations corresponding to higher tumor grade</td>
</tr>
<tr>
<td>Adamson et al., 1993 (33)</td>
<td>Urinary tissue factor (~35 units)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>53 with Pca; mean age, 72 years</td>
<td>37 with BPH (mean age, 70 years); 20 healthy controls (mean age, 34 years)</td>
<td>30 of 53 [57 (42–70)%]</td>
<td>43 of 57 [75 (62–86)%]</td>
<td>Early morning midstream urine; centrifuged</td>
<td></td>
</tr>
<tr>
<td>Irani et al., 1997 (34)</td>
<td>Ratio of serum to urinary PSA* (1.25)</td>
<td>57 with Pca; median age, 73.5 years</td>
<td>73 with BPH (by biopsy); median age, 68 years</td>
<td>48 of 57 [84 (72–93)%]</td>
<td>65 of 73 [89 (82–96)%]</td>
<td>24-h urine; stored at ~20 °C</td>
<td></td>
</tr>
<tr>
<td>Tremblay et al., 1987 (35)</td>
<td>Urinary PSA (20 μg/24 h urine)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>13 with Pca; ages not mentioned</td>
<td>8 with BPH by biopsy (age range, 60–90 years); 18 healthy volunteers by DRE (age range, 20–45 years)</td>
<td>4 of 13 [31 (9–61)%]</td>
<td>26 of 26 [100 (87–100)%]</td>
<td>24-h urine; stored at 4 °C; 5-mL collections stored at ~80 °C</td>
<td></td>
</tr>
<tr>
<td>Irani et al., 2005 (36)</td>
<td>Ratio of urinary to serum PSA (not mentioned)</td>
<td>83 with Pca; median age, 70.1 years</td>
<td>82 controls (by biopsy); median age, 66.6 years</td>
<td>42% (see comment)</td>
<td>80% (see comment)</td>
<td>12-h urine; stored at ~70 °C</td>
<td>Sensitivity and specificity estimated from ROC curve; multicenter study; blinded analysis</td>
</tr>
</tbody>
</table>

<sup>a</sup>Percentage values in parentheses are the 95% confidence interval.
<sup>b</sup>TRAP, telomeric repeat amplification protocol; Pca, prostate cancer; UW, urethral washing; BPH, benign prostatic hyperplasia; MCM-5, minichromosome maintenance 5; BHUAE, basic human arginine amidase.
<sup>c</sup>Estimated from figure.
<sup>d</sup>95th centile of range of values for controls.
<sup>e</sup>Urinary PSA expressed as μg/mmol of creatinine.
<sup>f</sup>Urinary PSA expressed as ng/mL.
seemed to be no differences in the urine-handling procedures or in age between cases and controls in the 2 studies. The sensitivity and specificity for bladder tumor fibronectin were 43% and 77%, respectively (25), and for basic human arginine amidase were 47% and 100%, respectively (26).

Two studies assessed scatter factor (27, 28), 2 other studies assessed transferrin (or transferrin/creatinine ratio) (29, 30), and 1 study assessed immunoglobulin concentrations (31) as potential biomarkers. All 5 studies reported rather poor performance characteristics as these potential markers showed either a sensitivity (27–29, 31) or a specificity (30) near 30%. Furthermore, 3 of the studies (27, 28, 30) did not mention the ages of the participants, and 1 study (31) might have used a different sampling for cases and controls. In addition, the females used as controls by Rosen et al. (28) could not be excluded from specificity calculations.

Tissue factor may be expressed by malignant tissue and aid tumor growth (32). Both Lwaleed et al. (32) in 2000 and Adamson et al. (33) in 1993 obtained quite similar results for urinary tissue factor, with sensitivities of 57% (33) and 65% (32) and specificities in both studies around 75%. Furthermore, Lwaleed et al. (32) reported a significant increase in urinary tissue factor concentrations with higher tumor grade. However, both studies (32, 33) included large proportions of rather young persons as controls. Lwaleed et al. (32) also used females as controls.

Three studies evaluated urinary PSA, which may originate from free serum PSA or be produced in the urethral duct (34), as a tumor marker in 24-h urines (34, 35) and 12-h urines (36), respectively. Tremblay et al. (35) obtained a sensitivity of 31% and a specificity of 100%, but the majority of the controls were much younger than the cases. Age differences between cases and controls were small in the studies by Irani and coworkers. (34, 36), which examined the ratio between urinary and serum PSA. The first study showed good performance, with 84% sensitivity and 89% specificity (34), but these results were not confirmed in a later multicenter study with more participants (sensitivity, 42%; specificity, 80%) (36).

Studies evaluating urine tests based on qualitative protein markers are listed in Table 5. Each of the 4 markers was evaluated in just 1 study. The best performance characteristics of all studies included in this review were reported by Edward et al. (37) in 1982 for prostatic cancer antigen 1, but the study population was rather small and the results have not been reproduced by any subsequent published studies in the last 23 years. In their rather large study, Chopin et al. (38) collected 24-h urines but obtained a very poor sensitivity (18%) for acidic fibroblast growth factor. Much higher sensitivities but somewhat lower specificities were reported in blinded studies by Rogers et al. (39) for α-methylacetyl-coenzyme A racemase and by Moses et al. (40) for matrix metalloproteinases. In the latter study, however, there were consid-
erable differences concerning age, sex, and collection of urine samples between cases and controls.

Discussion

Our literature search identified 34 studies evaluating 21 different urinary markers, including 2 genetic markers, 1 epigenetic marker, 3 RNA-based markers, 11 quantitative protein markers, and 4 qualitative protein markers. Sample sizes for most studies were small, and the majority of markers were evaluated in just 1 study each. Sensitivity and specificity varied widely. Because study populations were mostly preselected according to suspicion of prostate cancer (a prerequisite for prostate biopsy), the observed sensitivities and specificities typically reflect the diagnostic value over and above the commonly used primary screening tests rather than the diagnostic value in a primary screening situation. In these studies, suspicion of prostate cancer frequently was based on serum PSA measurements or DRE (pertinent details were given in some, but not all studies). For the same reason, the reported sensitivities and specificities of urine-based markers reported in this study may not be comparable to estimates of sensitivity and specificity for serum PSA concentrations obtained in different settings, such as the large prospective evaluation by Gann et al. (41), who reported a sensitivity and specificity of 46% and 91%, respectively. Nevertheless, if the high sensitivities and specificities reported for some urine markers could be replicated and estimated with higher precision in future, prospectively designed studies, there might be potential to enhance screening for prostate cancer by use of urinary markers, alone or in combination with serum PSA. In addition to performance characteristics, practical aspects must be considered in this context.

Particularly advantageous in terms of practicality of a screening test would be a stable marker that is not strongly influenced by temperature, so that urine samples could be mailed to laboratories for analysis. In most of the studies, the urine specimens were stored frozen, which might be difficult to achieve in mass screening programs. In nearly all of those studies, it was not clear whether and to what extent the stability of the markers would be affected by intermediate storage of samples at room temperature.

Another important practical aspect is the type of urine sample needed for analysis. For example, although in their first study Irani et al. (34) obtained promising results for urinary PSA with 24-h urine samples, in studies using midstream urine samples, results were disappointing (42, 43). Even in the later multicenter study by Irani et al. (36), in which they used 12-h urine samples, the good results from the previous study (34) were not reproduced. Whether the different procedures used for urine collection were the only cause of the discrepancy remains to be examined. It is evident, however, that a marker that could be measured in more easily collected samples (such as a midstream urine sample compared with a 24-h urine) would be a great advantage for use in mass screening.

In 9 of 13 studies evaluating DNA- or RNA-based markers, prostatic massage or palpation was performed with the idea of increasing the sensitivity. This was not done in any of the studies evaluating protein-based markers. When we compared the performance of glutathione S-transferase P1 in the 2 studies from Goessl and coworkers (10, 11) (with massage) with the performance in the study by Jeronimo et al. (12) (without massage), prostatic massage seemed to be associated with a much higher sensitivity. However, in the recent study by Crocitto et al. (14), both the sensitivity and specificity were poor despite prostatic massage. The impact of prostatic massage has not been evaluated within a single study using otherwise consistent methodology among men with and without prostatic massage; therefore, its impact remains unclear. Clarification of this issue appears to be important because of the possible lack of acceptance by patients and the additional work required by physicians when prostatic massage would need to be performed as part of a screening test. A similar problem of acceptance could occur when urine is collected by urethral milking/washing or by catheter, as was done in 6 studies. In contrast to prostatic massage, there was no indication that the test performance might be increased by these methods.

For DNA- and RNA-based markers, more extensive laboratory processing may often be required than for protein-based markers. Compared with RNA, DNA has the advantage of being mostly more stable in urine; therefore, DNA-based markers might require much less effort for preservation of urine samples. Immunologic assays such as ELISA dominate the analyses of protein markers. Because these tests can be quite inexpensive, a stable and reliable protein-based marker would appear to be particularly suited for mass screening. New mass spectrometric techniques, such as matrix-assisted laser desorption/ionization time of flight (MALDI-TOF) or surface-enhanced laser desorption/ionization-time of flight (SELDI-TOF) mass spectrometry, might open further avenues for protein-based screening for prostate cancer in the future. Rehman et al. (44) used this method to identify proteins in urine samples as markers for prostate cancer. Their results appeared to be potentially promising, but the study had to be excluded from this review because there were only 6 cases and 6 controls.

Apart from the small sample size leading to rather imprecise estimates of performance characteristics of tests in most studies, potential sources of bias also must be taken into account. If possible, evaluation of diagnostic tests should be performed in a blinded fashion. Only 7 studies explicitly reported using a blinded design (13, 15, 19, 24, 36, 39, 40). One of these studies (40) used different urine sampling methods among cases and controls, which may hinder comparability of results. In 3 studies (28, 32, 40), the specificity may have been overestimated by inclusion of female participants as controls.
because it was not possible to exclude them from the specificity calculations. In addition, bias caused by age differences between cases and controls could not be excluded in 18 studies and should be carefully avoided in future studies.

In 18 studies, at least some information about tumor stage, grade, or Gleason score was provided, but only for 3 studies were the performance characteristics calculated according to stage or Gleason score (10, 19, 31), and only 5 other studies provided some information on the association between marker and tumor stage or Gleason score (11–13, 18, 32). Therefore, for most markers the sensitivity for detecting early-stage prostate cancer, the main target of potential screening programs, is unknown. Likewise, the ability of tests to distinguish aggressive from slowly growing cancers is essentially unknown. Future studies should aim at differentiating estimates of sensitivity according to stage and grade.

Because of the differences in study populations, collection and handling of urine samples, and laboratory techniques, the comparability of estimates of performance characteristics between studies is limited. For the same reasons, and because few markers were evaluated by more than 1 study, we decided not to use formal meta-analysis techniques to pool results from multiple studies.

The fact that some of the most favorable results, such as those reported by Edwards et al. in 1982 (37), have never been replicated also points to potential publication bias, which should be kept in mind when interpreting the summary tables provided in this review.

In summary, development of urinary biomarkers for detection of prostate cancer appears to be in an early phase. On the basis of the 5-phase model of biomarker development for early detection of cancer, proposed by Sullivan Pepe et al. (45), the studies evaluated in this review appear to be phase 1 or 2. Some of the studies reported promising results that should be replicated in larger, prospectively designed studies. Progress in molecular biology might offer new opportunities to identify urine markers with better sensitivity and specificity. Mass spectrometry might be a particularly promising approach for discovering novel markers, but it should be kept in mind that peak height, which is measured by the mass spectrometer, is not linearly related to the abundance of the specific molecule and that the peaks identified by different investigators as indicative of the same disease are different (46). Furthermore, a combination of markers might be useful to improve the performance characteristics. Ideally, future prospectively designed studies should assess a variety of markers in the same, large study population to provide comparable, precise estimates of sensitivity and specificity for each individual marker as well as for various combinations of markers. In addition, future studies should also address practical issues, such as urine sample collection procedures or the need to immediately freeze samples at subzero temperatures, which might be relevant for mass screening. Finally, the costs of marker analysis and the impact of their possible use on the cost-effectiveness of population-based mass screenings also need to be addressed.

The work of Heiko Müller was supported in part by a grant from the German Ministry for Family Affairs, Senior Citizens, Women and Youth.

References


