Comparison of the diagnostic performance of fluorescence in situ hybridization (FISH), nuclear matrix protein 22 (NMP22), and their combination model in bladder carcinoma detection: a systematic review and meta-analysis

Qindong Liang¹
Guangjie Zhang²
Wuxian Li³
Jing Wang⁴
Shangchun Sheng¹

¹Department of Clinical Laboratory, Affiliated Hospital & Clinical Medical College of Chengdu University, Chengdu, Sichuan, China; ²Department of Clinical Laboratory, Chengdu Fifth People’s Hospital, Chengdu, Sichuan, China; ³Department of Clinical Laboratory, Chongqing Health Center for Women and Children, Chongqing, China; ⁴Department of Blood Transfusion, The First Affiliated Hospital of Chongqing Medical University, Chongqing, China

Background: Emerging studies reported that combination of fluorescence in situ hybridization (FISH) and nuclear matrix protein 22 (NMP22) could increase the sensitivity and specificity of bladder carcinoma (BC) management. Nevertheless, the reports remain inconsistent. This meta-analysis was undertaken to evaluate the diagnostic performance of FISH, NMP22, and their combination model in BC.

Materials and methods: A systematic literature search was carried out in PubMed, Embase, Cochrane Library, Web of Science, Chinese National Knowledge Infrastructure, and Wanfang database dated up to October 2018. Suitable studies were identified and raw data were extracted. Meta-analysis was conducted to calculate the global sensitivities, specificities, likelihood ratio, diagnostic odds ratio (DOR), and the areas under the summary receiver operating characteristic (SROC) curves for FISH, NMP22, and their combination model, separately. All the meta-analysis estimates were derived using STATA (version 12.0) and MetaDisc (version 1.4) software packages.

Results: Seven eligible studies were included for analysis. The global sensitivities with 95% CI for FISH, NMP22, and their combination model were 0.79 (95% CI: 0.75–0.83), 0.76 (95% CI: 0.71–0.81), and 0.82 (95% CI: 0.75–0.88); specificities were 0.85 (95% CI: 0.76–0.91), 0.70 (95% CI: 0.55–0.81), and 0.90 (95% CI: 0.70–0.97); DORs were 22.215 (95% CI: 10.695–46.144), 7.365 (95% CI: 3.986–13.610), and 41.940 (95% CI: 13.546–129.853); and the areas under the SROC curves were 0.86 (95% CI: 0.82–0.88), 0.79 (95% CI: 0.76–0.83), and 0.90 (95% CI: 0.87–0.92).

Conclusion: Our systematic review implied that the diagnostic performance of combination model of FISH plus NMP22 may outperform FISH or NMP22 alone in BC detection.

Keywords: bladder cancer, fluorescence in situ hybridization, nuclear matrix protein 22, diagnosis, meta-analysis

Introduction

Bladder carcinoma (BC), affecting the genitourinary system with raising morbidity and mortality, is one of the most prevalent urological malignancies worldwide.¹ Nowadays, cystoscopy still remains the gold standard in the diagnosis and surveillance of BC.² However, the uncomfortable experience for patients and high costs confine the clinical
utility of cystoscopy. Thereby, alternative noninvasive diagnostic methods with lower cost have been sought.

Fortunately, several urine-based tumor biomarkers, like fluorescence in situ hybridization (FISH), nuclear matrix protein 22 (NMP22), ImmunoCyt (uCyt+), and urinary cytology (Cyt), have been developed and approved by the US Food and Drug Administration for the BC diagnosis or surveillance. Among them, FISH test is used to identify BC-related genetic alterations based on several probes to chromosomes 3, 7, 17, and 9p21. NMP22 plays a significant role in the construction of the nuclear framework and takes participant in some biological progresses. The NMP22 test is an enzyme immunoassay, which has been used to detect the amount of NMP22 in voided urine. uCyt+ identifies exfoliated BC cells using fluorescence-labeled monoclonal antibodies. Similarly, Cyt, which is the most widely accepted noninvasive biomarker for BC diagnosis and prognosis, detects free BC cells in urine based upon Papanicolaou or Wright stain. However, unfortunately, none of them got a sensitivity (SEN) and specificity (SPE) high enough as diagnostic biomarkers in BC detection.

Interestingly, emerging studies reported that combination of these tumor biomarkers could increase the SEN and SPE of BC management. Especially, the diagnostic performance of the combination model of FISH and NMP22 is higher than FISH or NMP22 alone. Nevertheless, the results remain controversial. To summarize and compare the diagnostic values of FISH, NMP22, and their combination model, this systematic review and meta-analysis was conducted depending on seven available studies from several online databases.

Materials and methods
This systematic review was carried out based on the Preferred Reporting Items for Systematic reviews and Meta-analyses statement. Literature search strategy
A computerized literature search was conducted based on PubMed, Embase, Cochrane Library, Web of Science, China National Knowledge Infrastructure (CNKI), and Wanfang Database until October 2018. The search terms used for literature retrieval included 1) “bladder cancer” or “bladder carcinoma” or “bladder tumor;” 2) “fluorescence in situ hybridization” or “FISH;” 3) “nuclear matrix protein 22” or “NMP22;” 4) “diagnostic” or “diagnosis.” In addition, the references cited by all the eligible literatures were hand-searched for relevant citations. No language restrictions were used in this meta-analysis.

Study selection
Two authors (Qindong Liang and Guangjie Zhang) independently reviewed all the searched articles and selected the citations suitable for the meta-analysis. Studies should be enrolled if they met all the following criteria: 1) patients with BC were confirmed by cystoscopy or pathological examination; 2) samples were taken before any manipulation; 3) the number of patients or control groups was more than 20; 4) the significance of FISH, NMP22, and their combination model in bladder cancer detection were evaluated; and 5) available data to create 2 by 2 contingency tables. On the other hand, letters, abstracts, case reports, meetings, and reviews were removed. For duplicate studies, we enrolled the publication with the largest number of participants or the one with most informative data.

Data extraction and quality assessment
The same two authors scrutinized all the full-text of included studies, and the eligible raw data of each enrolled studies were extracted, including the first author’s name, year of publication, country, the number of cases and control groups, study design characteristics, the diagnostic information (such as SEN, SPE, cutoff value, etc). The exact number of 2 by 2 contingency tables was extracted directly or recalculated.

According to the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) checklist, the risk of bias of included studies was assessed. The QUADAS-2, which is recommended for the quality assessment of diagnostic investigation, is a reformativ tool containing four vital domains: 1) patient selection; 2) index test; 3) reference standard; and 4) flow and timing. Each domain includes two or three signal questions to aid judging the risk of bias. Using the QUADAS-2, the risk of bias for each domain was rated as “high,” “low,” and “unclear.” Any discrepancy was solved in consensus meetings, and if necessary, arbitration by another investigator (Wuxian Li).

Statistical analysis
STATA 12.0 (http://www.stata.com; StataCorp, College Station, TX, USA) and Meta-Diac 1.4 (http://www.hrc.es/investigacion/metadisc_en.htm); Unit of Clinical Biostatistics, RamOn y Cajal Hospital, Madrid, Spain) statistical software packages were adopted to perform all the statistical analyses. The heterogeneity between the eligible studies was evaluated by the standard chi-squared statistic and the inconsistency index (I²). The global SEN, SPE, positive likelihood ratio (PLR), negative likelihood ratio (NLR), and diagnostic odds ratio (DOR) with the corresponding 95% CI...
for FISH, NMP22, and their combination were computed. Simultaneously, area under the summary receiver operating characteristic (SROC) curve was calculated to quantify the diagnostic power. Furthermore, meta-regression was conducted to explore the sources of heterogeneity. Finally, publication bias was evaluated by Deeks’ and Begg’s regression tests, and the P-value below 0.05 was thought to have statistical significance.

Results
Study selection
The computerized search yielded a total of 169 records: 28 for PubMed, 95 for Embase, one for Cochrane Library, 31 for Web of Science, and seven for CNKI and Wanfang Database, separately. According to the flowchart (Figure 1), 37 duplicate records were removed. Furthermore, after reading the titles and abstracts, 111 studies were eliminated as they were irrelevant studies, abstracts, books, letters, meetings, and comments. Consequently, there were 21 studies left for full-text reading. After scrutinizing the full-text of the 21 studies, 14 studies were discarded: two did not evaluate diagnostic values of FISH or NMP22,\textsuperscript{26} one with number of controls <20,\textsuperscript{10} three were prognostic analyses,\textsuperscript{11,27,28} and eight studies did not report diagnostic performance of the combination of FISH and NMP22.\textsuperscript{12–14,29–33} Finally, seven studies were suitable for our meta-analysis.\textsuperscript{16–22}

Study characteristics
Among the seven eligible studies, four studies were from China, two from Germany, and one from Korea. All the included studies were published between 2009 and 2016. Barring one study, FISH and NMP22 detection had been successfully performed in all the participants of the included studies. In this study,\textsuperscript{18} FISH and NMP22 BladderCheck were conducted in 119 and 149 participants, respectively (reasons not shown). Finally, a total of 1,596, 1,566, and 1,549 individuals for FISH, NMP22, and their combination in seven eligible studies were recruited in this meta-analysis, separately. The main characteristics of the seven included studies, including name of the first author, year of publication, country, number of cases and controls, SEN and SPE, are listed in Table 1.

Quality assessment
Based on the QUADAS-2, the quality assessment of the seven included studies was demonstrated by a picture of “risk of bias” and “applicability concerns” (Figure 2). The major bias within the seven studies was focused upon the “patient selection” and “reference standard.” Especially, only two studies fulfilled all the questions of each domain;\textsuperscript{19,22} three studies did not state whether the reference standard results were interpreted without the knowledge of the results of the index test;\textsuperscript{18,20,21} moreover, in one study, not all the patients were included in FISH or BladderCheck tests.\textsuperscript{18} In general, all the seven eligible studies were of upper middle quality.

Heterogeneity exploration and meta-regression
Spearman test results indicated that there was no heterogeneity between the included studies caused by threshold effects (Table 2). However, heterogeneity originated from other factors except threshold effects for FISH, NMP22,
and combination model of FISH and NMP22 (COMB) was found in our meta-analysis. In order to explore the sources of heterogeneity, meta-regression was conducted using several covariates (such as ethnicity, assay kit, and mean age). According to the meta-regression analysis (calculated with MetaDisc 1.4), we found that for FISH, the relative diagnostic odds ratio (RDOR) of ethnicity was 2.62, and the \( P \) value was 0.0421 (<0.05), which implied that the ethnicity was probably the source of heterogeneity for FISH in BC diagnosis.

### Diagnostic accuracy analysis

All the diagnostic values of FISH, NMP22, and their combination model are shown in Figures 3–5. The global SEN of FISH was 0.79 (95% CI: 0.75–0.83); SPE was 0.85 (95% CI: 0.76–0.91); PLR was 5.410 (95% CI: 3.201–9.145); NLR was 0.244 (95% CI: 0.189–0.313); and DOR was 22.215 (95% CI: 10.695–46.144). The area under the SROC curve was 0.86 (95% CI: 0.82–0.88).

### Table 1 Characteristics of the seven studies included in this meta-analysis

<table>
<thead>
<tr>
<th>Author (year)</th>
<th>Country</th>
<th>BC/controls</th>
<th>Grade</th>
<th>Stage</th>
<th>Mean age (year)</th>
<th>Assay kit</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Horstmann (2009)(^6)</td>
<td>Germany</td>
<td>113/108</td>
<td>G(_1)/G(_2)/G(_3) (3/25/33)</td>
<td>NR</td>
<td>79.0</td>
<td>Abbott</td>
<td>76%</td>
<td>68%</td>
</tr>
<tr>
<td>Kim (2009)(^7)</td>
<td>Korea</td>
<td>106/50</td>
<td>Low/high (5/4/52)</td>
<td>pT(_1)/pT(_2)-4 (64/42)</td>
<td>62.9</td>
<td>Abbott</td>
<td>84.9%</td>
<td>75.5%</td>
</tr>
<tr>
<td>Li (2013)(^8)</td>
<td>China</td>
<td>104/71</td>
<td>G(_1)/G(_2)/G(_3) (29/18/18)</td>
<td>NR</td>
<td>62.4</td>
<td>Abbott</td>
<td>86.5%</td>
<td>67.6%</td>
</tr>
<tr>
<td>Todenhöfer (2013)(^9)</td>
<td>Germany</td>
<td>115/693</td>
<td>G(_1), G(_2) + CIS (82/44)</td>
<td>NR</td>
<td>67</td>
<td>Abbott</td>
<td>71.3%</td>
<td>84.4%</td>
</tr>
<tr>
<td>Ran (2014)(^10)</td>
<td>China</td>
<td>43/25</td>
<td>Low/high (19/24)</td>
<td>NMIBC/MIBC (25/18)</td>
<td>63.7</td>
<td>Linked-Biotech pathology</td>
<td>81.4%</td>
<td>86%</td>
</tr>
<tr>
<td>Jiang (2015)(^11)</td>
<td>China</td>
<td>69/31</td>
<td>NR</td>
<td>NR</td>
<td>60.5</td>
<td>Matritech</td>
<td>75.4%</td>
<td>72.5%</td>
</tr>
<tr>
<td>Pang (2016)(^12)</td>
<td>China</td>
<td>71/29</td>
<td>G(_1)/G(_2)/G(_3) (24/31/16)</td>
<td>NR</td>
<td>62.16</td>
<td>Matritech</td>
<td>78.87%</td>
<td>80.28%</td>
</tr>
</tbody>
</table>

**Abbreviations:** BC, bladder cancer; NR, not report; FISH, fluorescence in situ hybridization; NMP22, nuclear matrix protein 22; COMB, combination model of FISH and NMP22.
Table 2 Spearman test, Deeks’ regression, and Beeg’s regression results for FISH, NMP22, and COMB

<table>
<thead>
<tr>
<th></th>
<th>Spearman test</th>
<th>Deeks’ regression</th>
<th>Beeg’s regression</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>P-value</td>
<td>t</td>
</tr>
<tr>
<td>FISH</td>
<td>-0.500</td>
<td>0.253</td>
<td>0.73</td>
</tr>
<tr>
<td>NMP22</td>
<td>0.536</td>
<td>0.215</td>
<td>2.43</td>
</tr>
<tr>
<td>COMB</td>
<td>0.500</td>
<td>0.253</td>
<td>1.77</td>
</tr>
</tbody>
</table>

Abbreviations: FISH, fluorescence in situ hybridization; NMP22, nuclear matrix protein 22; COMB, combination model of FISH and NMP22.

Figure 3 Forest plots for sensitivity and specificity of FISH (A), NMP22 (B), and combination model of FISH and NMP22 (C) in BC diagnosis.

Abbreviations: FISH, fluorescence in situ hybridization; NMP22, nuclear matrix protein 22; BC, bladder carcinoma.
The global SEN of NMP22 was 0.76 (95% CI: 0.71–0.81); SPE was 0.70 (95% CI: 0.55–0.81); PLR was 2.517 (95% CI: 1.643–3.856); NLR was 0.342 (95% CI: 0.268–0.436); and DOR was 7.365 (95% CI: 3.986–13.610). The area under the SROC curve was 0.79 (95% CI: 0.76–0.83).

**Combination model of FISH and NMP22**

The global SEN of combination of FISH and NMP22 (COMB) was 0.82 (95% CI: 0.75–0.88); SPE was 0.90 (95% CI: 0.70–0.97); PLR was 8.166 (95% CI: 2.603–25.616); NLR was 0.195 (95% CI: 0.139–0.273); and DOR was 41.940 (95% CI: 13.546–129.853). The area under the SROC curve was 0.90 (95% CI: 0.87–0.92).

**Publication bias**

No evidence of obvious publication bias was found by Deeks’ regression test of asymmetry in the analysis of either FISH or NMP22 or COMB (Table 2 and Figure 6). Furthermore,
Comparison of FISH, NMP22, and their combination in bladder carcinoma

Figure 5 Fagan's nomogram of FISH (A), NMP22 (B), and combination model of FISH and NMP22 (C) in BC diagnosis. Abbreviations: FISH, fluorescence in situ hybridization; NMP22, nuclear matrix protein 22; BC, bladder carcinoma; LR, likelihood ratio.

Figure 6 Deeks' regression results of publication bias of FISH (A), NMP22 (B), and combination model of FISH and NMP22 (C). Abbreviations: FISH, fluorescence in situ hybridization; NMP22, nuclear matrix protein 22; ESS, effective sample size.
the Beeg’s test results also validate the aforementioned results (Table 2).

Discussion

Although cystoscopy is currently a mainstay for BC diagnosis, the high cost for a cystoscopy and the risk of complications limit the full clinical utility. Therefore, the noninvasive and low-cost diagnosis of BC remains challenging. Nowadays, cytology and several urine-based biomarkers, like FISH, NMP22, and bladder cancer antigen, show importance for BC diagnosis and surveillance. The diagnostic performance as single tumor markers have been described widely,\(^8,18,34–37\) However, none of these biomarkers are recommended for the routine application owing to the low SEN or SPE.\(^38–40\) Surprisingly, combination of these urine-based biomarkers could provide satisfactory diagnostic values for BC.\(^10,16,17,19–22\)

To the best of our knowledge, this systematic review and meta-analysis is the first to compare the diagnostic performance of combination model of FISH and NMP22 (COMB) with FISH alone or NMP22 alone in BC detection. In our study, we identified seven eligible studies with 1,596, 1,566, and 1,549 participants for FISH, NMP22, and their combination model, respectively. The global SEN of COMB for BC is 0.82 slightly higher than FISH and NMP22 alone (0.79 and 0.76), and the SPE of COMB is 0.90 also higher than FISH and NMP22 alone (0.85 and 0.70). This indicates that the combination model of FISH and NMP22 may present a higher SEN and SPE than FISH and NMP22 alone in BC diagnosis. Area under the curve (AUC) is an insightful indicator to estimate the overall diagnostic performance. In our study, the AUC of the SROC curve for COMB in BC detection is 0.90, compared with FISH and NMP22 alone (0.86 and 0.79), which means a good diagnostic accuracy of COMB. DOR is another single metric of diagnostic accuracy. A higher DOR means a better discriminating validity. The global DOR for COMB is 41.94 which is much higher than FISH and NMP22 alone (22.215 and 7.365). In general, the diagnostic performance of the combination model of FISH and NMP22 outperforms FISH and NMP22 alone, and the combination of FISH and NMP22 may be a new model in BC diagnosis.

As shown in Figure 3, substantial heterogeneity exists between the included studies for FISH and NMP22 as single biomarker and their combination model in BC diagnosis (overall F for FISH is 81.91%, NMP22 95.51%, and combination model 98.42%). Spearman tests show a \(P\)-value of 0.253 for FISH, 0.215 for NMP22, and 0.253 for COMB, separately, indicating no evidence of threshold effects. Therefore, meta-regression was conducted to further explore the source(s) of heterogeneity. Consequently, ethnicity was found to be the main source of heterogeneity for FISH, NMP22, and their combination model (Table 2).

Our findings on diagnostic accuracy of FISH and NMP22 in BC detection were in accordance with those of previous meta-analyses.\(^41,42\) However, our study has some advantages. First, no language restrictions were applied in this meta-analysis and more eligible studies beyond English version were enrolled. Second, we are the first to pool the diagnostic values of the combination model of urine-based tumor markers (FISH and NMP22).

Our study has several limitations. First, among the seven included studies, FISH and NMP22 BladderChek tests were not successfully performed in all the participants.\(^18\) Second, the eligible studies in our study differ in some ways, such as methodological quality, mean age, assay kit, and ethnicity. Obvious heterogeneity was observed between the included studies. After meta-analysis, we found that different ethnicity may be the main source of heterogeneity. Third, the number of the included studies is small. We have conducted a comprehensive and meticulous search in PubMed, Embase, Cochrane Library, Web of Science, CNKI, and Wanfang database. After reviewing the titles, abstracts, and full-text of studies eligible for our study, only seven studies met the requirements and were included. Finally, as positive results are easy to be published, potential publication bias may still exist, even though there is no significant publication bias presented by Deeks’ and Beeg’s regression analyses.

Conclusion

Our study demonstrates that the combination model of FISH and NMP22 has a satisfactory diagnostic performance outperforming FISH alone and NMP22 alone in BC detection. However, more multicenter, well-designed, prospective trials should be carried out to verify the diagnostic accuracy of the combination model of FISH and NMP22 in BC detection.

Acknowledgment

This work was supported by the youth science fund project of National Natural Science Foundation of China (NO 81702101).

Author contributions

All the authors mentioned in this article have contributed to this work. Shangchun Sheng conceptualized the study; Qindong Liang and Guangjie Zhang performed article retrieval; Qindong Liang, Guangjie Zhang, Wuxian Li, and
Jing Wang conducted data extraction and statistical analysis. Qindong Liang and Guangjie Zhang drafted the original manuscript. Shangchun Sheng, Wuxian Li, and Jing Wang critically revised the manuscript. All authors contributed to data analysis, drafting and revising the article, gave final approval of the version to be published, and agree to be accountable for all aspects of the work.

Disclosure
The authors report no conflicts of interest in this work.

References


