Sensitivity and Specificity of Diffusion-Weighted Magnetic Resonance Imaging in Diagnosis of Bladder Cancers

Abstract

Purpose: Our study is designed to examine the diagnostic performance of diffusion-weighted magnetic resonance imaging (DW-MRI) for bladder cancers (BC), and to determine whether DW-MRI can differentiate muscle invasive bladder cancer (MIBC) from non-MIBC (NMIBC).

Methods: A meta-analysis was performed of published studies that investigated the performance of DW-MRI for BC. These studies were retrieved from scientific literature databases using sensitive electronic search strategies. The STATA 12.0 and Meta-disc software were employed for statistical analyses of data extracted from selected studies.

Results: Our search initially returned 230 articles, of which 11 met the inclusion criteria and were enrolled into the final meta-analysis. Five of the included studies reported the diagnostic performance of DW-MRI for BC with a cumulative total of 243 BC patients and 82 healthy subjects. Eight studies investigated the diagnostic performance of DW-MRI for differentiating MIBC from NMIBC, involving 259 MIBC lesions and 515 NMIBC lesions. Meta-analysis results were as follows: the diagnostic performance of DW-MRI for BC (sensitivity: 0.95 [0.75-0.99]; specificity: 0.85 [0.74-0.92]; positive likelihood ratio: 6.45 [3.64-11.42]; negative likelihood ratio: 0.055 [0.009-0.333]; diagnostic odds ratio: 117.11 [19.37-708.05]; area under the curve (AUC): 0.91); the diagnostic performance of DW-MRI to differentiate MIBC from NMIBC (sensitivity: 0.85 [0.76 - 0.91]; specificity: 0.90 [0.87 - 0.93]; positive likelihood ratio: 8.81 [6.43 - 12.07]; negative likelihood ratio: 0.16 [0.10 - 0.28]; diagnostic odds ratio: 53.95 [25.68 - 113.33]; AUC: 0.92).

Conclusion: DW-MRI has an outstanding diagnostic performance, with advanced sensitivity and specificity, for imaging of bladder cancers and for differentiating MIBC from NMIBC.

Correspondence to:
Xian Li, MM, Department of Radiology, Jining Medical College, No.16, Hehua Road, Taibai Lake District, Jining 272067, P.R. China.
E-mail: lixian740@163.com;luoyahong925@126.com
Bladder cancer (BC) is the most frequently diagnosed urinary tract malignancy and accounts for 10% of all malignancies in males throughout the world [1,2]. BC ranks highest in the incidence of genitourinary cancers in China and, among men, ranks second only to prostatic cancer in the western world [3]. BC is a complex and multifactorial disease associated with both genetic susceptibility and environmental exposure. Smoking and occupational exposure to aromatic amines are two major risk factors for BC [4]. Among newly diagnosed cases of BC, approximately 70% are diagnosed with non-muscle invasive bladder cancer (NMIBC), while the remaining 30% exhibit muscle invasive bladder cancers (MIBC) [5]. Based on the American Joint Committee on Cancer and the International Union Against Cancer, NMIBC was seen in low-grade and stage T1 cancers, and MIBC was in high-grade and stage T2~T4 cancers [6]. Up to 70% of the patients with NMIBC still develop at least one recurrence within five years after transurethral resection and approximately 13% of these patients progress to MIBC or a more advanced disease [1,7]. Early diagnosis of BC is possible by detection of blood in the urine, however, accurate preoperative staging and disease prognosis is fundamental to prevent cancer recurrence and improve the overall success of treatment. Thus, it is essential to explore more effective tools to diagnose BC early and accurately, and to guide treatment strategies and clinical monitoring of BC patients [8].

In recent years, several MRI methods, in conjunction with conventional methods including computed tomography (CT), ultrasound and positron emission tomography (PET), are being used for routine radiodiagnostic applications [9]. Diffusion-weighted magnetic resonance imaging (DW-MRI) is a recent advancement in microstructural imaging modality, and is characterized within the context of functional MRI. DW-MRI allows tissue characterization based on the measurements of random Brownian motion of water molecules as a surrogate marker to obtain information on microstructural density of the imaged tissue [10]. Thus, accelerated or decelerated microscopic movements of water molecules relay the critical tissue information such as cellularity, cell density, and cell membrane permeability, which differ between normal, benign and tumor tissues. The successfully clinical application of DW-MRI is based on the principle that, through random and diffusion-driven displacement by water molecules, live tissue structures can be probed at a microscopic scale, which is beyond the imaging resolution of other methods [11]. In multiple previous studies, DW-MRI showed excellent performance in accurate characterization of lesions, and the results correlated with histological grading, clinical and pathological staging, differentiating tumour recurrence and guiding correct treatment decision in BC, and the staging accuracy was 96%, with a high diagnosis accuracy of 93.65% [9,12-15]. Nevertheless, other studies showed that there is not enough evidence to suggest a high diagnostic performance of DW-MRI, especially in early and accurate detection of BC [16]. In order to address this shortage of evidence, we conducted a meta-analysis by pooling high quality data from recent studies, published in medical journals, to examine the diagnostic performance of DW-MRI in BC and further determine whether DW-MRI could differentiate MIBC from NMIBC.

**Methods**

**Literature Search**

A computer-assisted literature search was conducted using the following databases: PubMed, Embase, EBSCO, Ovid, Springerlink, Wiley, Web of Science, Cochrane Library, Wanfang database, VIP and China National Knowledge Infrastructure, to identify relevant studies published prior to September 2014. To optimize our search strategy and to increase the chance of retrieving all eligible studies, the combination of keywords and search terms employed, including (“urinary bladder neoplasms” or “bladder neoplasm” or “bladder tumors” or “urinary bladder cancer” or “malignant tumor of urinary bladder” or “cancer of the bladder” or “bladder cancer” or “cancer of bladder”) and (“diffusion magnetic resonance imaging” or “diffusion MRI” or “diffusion weighted MRI”or “diffusion magnetic resonance imaging” or “DWI” or “diffusion-weighted magnetic resonance imaging” or “MRI-DWI” or “diffusion-weighted imaging” or “diffusion-weighted-MRI”).

All references of studies selected for the meta-analysis were examined; including a manual search to cross-check the completeness of the electronic searches.

**Eligibility Criteria**

Articles were eligible for inclusion into the meta-analysis if they met the following predetermined inclusion criteria: (1) diagnostic study; (2) BC patients diagnosed based on ESMO Practice Guidelines for diagnosis, treatment and follow-up [17]; (3) cystoscope biopsy as gold standard to determinate the degree of BC malignancy; (4) complete data related to sensitivity, specificity, positive likelihood ratio, negative likelihood ratio in four-by-fold table; (5) study published in English or Chinese. Of the retrieved articles, if the same
clinical and experimental trials were utilized by more than one study, we selected the latest article or the article that included the largest sample size. Simultaneously, we excluded studies with incomplete data and studies in which there was a considerable difference in baseline characteristics between cases and controls, duplicate studies or study subjects with unclear diagnoses.

**Data Items and Quality Control Assessment**

The following data items were extracted from each included full-text article and entered into a predetermined data-collection form independently by two investigators (Zhai and Li): article information (surname of first author, publication year, country and language), population characteristics (age, sex, ethnicity and number of incident cases), detection method, diagnostic method and number of lesions. Disagreements between the investigators during data extraction were resolved by a consensus meeting with a third reviewer (Wang), who is an expert in statistical analysis. Each of included full-text articles was scored using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool [18]:

- QUADAS01: Was there a wide spectrum of patients, including various disease spectrums?
- QUADAS02: Was the reference standard acceptable?
- QUADAS03: Was the time period between the reference standard and the detection test short enough?
- QUADAS04: Did the whole sample receive verification using a reference standard of diagnosis?
- QUADAS05: Did patients receive the same reference standard?
- QUADAS06: Was the reference standard independent of the index test?
- QUADAS07: Were the index test results judged without knowledge of the results of the reference standard?
- QUADAS08: Were the results of the reference standard judged without knowledge of the index test results?
- QUADAS09: Were the results obtained from the studies included in our meta-analysis consistent with the majority of clinical experimental results?
- QUADAS10: Were uninterpretable or intermediate test results reported?
- QUADAS11: Was an explanation provided for patients who did not complete the study?

**Statistical Methods**

Statistical analyses were performed with STATA 12.0 and Meta-disc to test the heterogeneity of the data and perform the current meta-analysis. Diagnostic odds ratio (DOR), specificity, sensitivity, positive likelihood ratio, negative likelihood ratio and area under the curve (AUC) in summary receiver operating curves (SROC) were calculated in either a fixed-effect model or random-effect model to evaluate the diagnosis of benign and malignant bladder tumours and MIBC by DW-MRI. Heterogeneity, including threshold effects and non-threshold effects, were assessed using the Spearman correlation coefficient between sensitivity and (1-specificity) logarithm by P value. P> 0.05 indicates the absence of threshold effects and P< 0.05 indicates the presence of threshold effects. Subsequently, the heterogeneity variation of non-threshold effects was evaluated using the P statistic. Higher values of I² suggested an increasing degree of heterogeneity. The random-effect model was applied if the significant heterogeneity existed (P<0.05 or I²>50%), otherwise the fixed-effect model was used [19]. If the heterogeneity caused by the existence of threshold effects was significant, the SROC was adopted to estimate the AUC as a measure of diagnostic accuracy. An AUC value approaching 1.0 indicated increasing diagnostic value. An AUC of 0.5-0.7 indicated poor discrimination and an AUC of 0.7-0.9 indicated moderate discrimination, whereas a value of greater than 0.9 suggested perfect discrimination. If the heterogeneity caused by the absence of threshold effects was significant, the parameters, including the sensitivity, the specificity, positive likelihood ratio, negative likelihood ratio, DOR and AUC, were pooled, analyzed and presented as a forest plot. Sensitivity analyses were carried out to evaluate the validity of the pooled outcomes by removal of individual study to observe the effect on the pooled outcomes. Fagan’s nomogram [20] was applied to assess the prior and posterior probability of the three diagnostic parameters with the Bivariate Boxplot to confirm whether the heterogeneity existed among studies.

Model diagnostics are not usually used in meta-analysis to reflect the results of sensitivity analyses, since meta-analysis is a data-processing procedure rather than a model fitting exercise [21]; however, as complex likelihood-based meta-analytic models have been applied, it is important to evaluate potential model misspecification as well as goodness of fit, and to identify outlying and influential data points. Midas, a comprehensive program of statistically graphical routines, was designed to evaluate diagnostic test performance in meta-analyses. It includes graphical models to check capabilities, including quantile plots of residual based...
The 515 NMIBC lesions. Seven studies were performed with Asian DW-MRI in 747 MIBC patients with 259 MIBC lesions and tissues with bladder lesions and 82 subjects with healthy bladder performance of DW-MRI in BC, including these 11 studies, five studies reported the diagnostic analysis [9,22-31]. Among these 11 studies, five studies reported the diagnostic performance of DW-MRI in BC, including 243 BC patients with bladder lesions and 82 subjects with healthy bladder tissues. Eight studies investigated the diagnostic performance of DW-MRI in 747 MIBC patients with 259 MIBC lesions and 515 NMIBC lesions. Seven studies were performed with Asian subjects and four with Caucasians (Table 1). The methodological quality assessment of each study on the basis of QUADAS is shown in Figure 1 and Figure 2. As shown in Figure 2, our included studies may need to be further studied in the following four areas: Was the delay between tests acceptable? Was incorporation avoided? Were uninterpretable results reported? Were subject withdrawals explained?

Results

Systematic Review

The electronic search followed by an extensive manual search originally retrieved a total of 230 studies. Of these, 27 studies were excluded after scanning their title and abstract and a further 192 studies were excluded for failing to meet our inclusion criteria after examining their full-text. Finally, 11 studies, with diagnostic trials published between 1998 and 2014, were enrolled into the final analysis [9,22-31]. Among these 11 studies, five studies reported the diagnostic performance of DW-MRI in BC, including 243 BC patients with bladder lesions and 82 subjects with healthy bladder tissues. Eight studies investigated the diagnostic performance of DW-MRI in 747 MIBC patients with 259 MIBC lesions and 515 NMIBC lesions. Seven studies were performed with Asian subjects and four with Caucasians (Table 1). The methodological quality assessment of each study on the basis of

<table>
<thead>
<tr>
<th>First author</th>
<th>Year</th>
<th>Country</th>
<th>Ethnicity</th>
<th>Language</th>
<th>Number</th>
<th>Gender (M/F)</th>
<th>Age (years)</th>
<th>MRI machine type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kobayashi S[22]</td>
<td>2011</td>
<td>Japan</td>
<td>Asians</td>
<td>English</td>
<td>104</td>
<td>38/66</td>
<td>81/38</td>
<td>Philips 1.5 T</td>
</tr>
<tr>
<td>Zhou GX[23]</td>
<td>2011</td>
<td>China</td>
<td>Asians</td>
<td>Chinese</td>
<td>71</td>
<td>25/46</td>
<td>30/40</td>
<td>Philips 3.0 T</td>
</tr>
<tr>
<td>Avcu S[9]</td>
<td>2011</td>
<td>Turkey</td>
<td>Caucasians</td>
<td>English</td>
<td>63</td>
<td>46/17</td>
<td>53/10</td>
<td>Siemens 1.5 T</td>
</tr>
<tr>
<td>Papalia R[25]</td>
<td>2012</td>
<td>Italy</td>
<td>Caucasians</td>
<td>English</td>
<td>36</td>
<td>34/38</td>
<td>25/15</td>
<td>GE 3.0 T</td>
</tr>
<tr>
<td>Zhang LH[26]</td>
<td>2012</td>
<td>China</td>
<td>Asians</td>
<td>Chinese</td>
<td>73</td>
<td>65/8</td>
<td>57/16</td>
<td>GE 3.0 T</td>
</tr>
<tr>
<td>Lista F[27]</td>
<td>2013</td>
<td>Spain</td>
<td>Caucasians</td>
<td>English</td>
<td>20</td>
<td>13/7</td>
<td>18/1</td>
<td>Siemens 1.5 T</td>
</tr>
<tr>
<td>Sevcenco S[29]</td>
<td>2014</td>
<td>Austria</td>
<td>Caucasians</td>
<td>English</td>
<td>43</td>
<td>10/33</td>
<td>NR</td>
<td>Siemens 3.0 T</td>
</tr>
</tbody>
</table>

M = male, F = female, MRI = magnetic resonance imaging, NR = not reported.
and negative likelihood ratio (NRP < 0.1) indicates the confirmation and exclusion of the diagnosis, the top right region (LRP > 10 and NRP > 0.1) indicates confirmation of the diagnosis, the bottom left region (LRP < 10 and NRP < 0.1) indicates exclusion of the diagnosis, and the bottom right region (LRP < 10 and NRP > 0.1) indicates that neither confirmation nor exclusion of diagnosis can be made. The result summarized from likelihood ratio scattergram suggested the limited clinically diagnostic significance of DW-MRI in BC (Figure 4C). The prior probability of 20%, the posterior probability of positive likelihood ratio of 62% and negative likelihood ratio of 1% related to the diagnostic performance of DW-MRI in BC are shown in the Fagan’s nomogram graph (Figure 4D).

Diagnostic Performance of DW-MRI to differentiate MIBC from NMIBC

The data extracted from eight studies were pooled and used to evaluate the diagnostic performance of DW-MRI in MIBC. Q test indicated the existence of significant heterogeneity ($P < 0.01, I^2 = 67.73\%$); as a result, the random-effect model was used during the evidence synthesis. The Spearman correlation coefficient between sensitivity and (1-specificity) logarithm was 0.071 with $P = 0.867$, which demonstrated the absence of threshold effect and the feasibility of pooled sensitivity,
specification, negative and positive likelihood ratios. Meta-analysis revealed that the sensitivity of the diagnostic performance of DW-MRI in MIBC was 0.85 (95%CI=0.76-0.91) (Figure 3D), specificity was 0.90 (95%CI=0.87-0.93) (Figure 3E), DOR was 53.95 (95%CI=25.68-113.33) (Figure 3F) and positive and negative likelihood ratio were 8.81 (95%CI= 6.43-12.07) and 0.16 (95%CI=0.10-0.28), respectively (Figure 4E and 4F). The result summarized from likelihood ratio scattergram suggested the limited clinically diagnostic significance of DW-MRI in distinguishing MIBC from NMIBC (Figure 4G). The prior probability of 20%, the posterior probability of positive likelihood ratio of 69% and negative likelihood ratio of 4% related to the diagnostic performance of DW-MRI in MIBC, are shown in Fagan’s nomogram graph (Figure 4H).

**Sensitivity Analyses**

As shown in Figure 5A, 5B, 5E and 5F, goodness of fit and bivariate normality test showed that the bivariate random-effects model could well fit the pool data. In addition,
FIGURE 5. The summary receiver operating curve and sensitivity analyses of the diagnostic performance of diffusion-weighted magnetic resonance imaging to differentiate muscle invasive bladder cancer from non-muscle invasive bladder cancer (A-D: BC patients vs. healthy subjects, E-H: MIBC vs. NMIBC).
influence analysis (Figure 5C, 5G) and outlier detection (Figure 5D, 5H) did not identify any biased studies. These analyses indicated that the pooled results were robust.

Discussion

BC is one of the most common malignancies of the urinary system with high recurrence rate and rapid progression. The treatment and prognosis of BC is dependent on its pathological phenotypes [12]; thus, accurate diagnosis of the aggressiveness of BC and precise prediction of the tumour behavior prevents over-diagnosis, under-diagnosis, under-treatment or over-treatment, leading to significant benefits to patient outcomes. In this study, a meta-analysis was performed to confirm the diagnostic performance of DW-MRI for BC based on all currently available published data. Our findings suggest that DW-MRI for BC shows outstanding diagnostic performance, with advanced sensitivity and specificity. As the method most frequently used to diagnose BC, PET has capacity of depicting key functional data about tumour metabolism. It has been widely applied in oncology: some types of malignant tumours show increased metabolic activity leading to increased glucose uptake, which can be imaged using 18F-fluorodeoxyglucose (18F-FDG); however, its application in the assessment of primary bladder cancer is limited because the urinary excretion of the tracer affects tumour visualization in the bladder [32]. DW-MRI provides precise information on cell density, tissue architecture and membrane integrity by measuring the diffusion of water molecule in vivo, with lower loss of signal representing low water diffusion and higher loss of signal representing high water diffusion [33,34]. Malignant tumours are characterized by active proliferation, enlarged nuclei, increased nuclear/cytoplasmic ratio, with ambiguous contour profiles in which many cells are tightly packed together with little extracellular space [31]. Consequently, malignant tumours exhibit high signal intensity by DW-MRI, attributed to their higher cellularity, increased extracellular tortuosity and tissue disorganization resulting in restricted water diffusion [35]. Yamada et al. reported that the addition of diffusion-weighting to conventional MRI may not only improve the accuracy of detection BC, but also may aid in the prognosis of BC [36].

Another finding of our study was that DW-MRI can differentiate MIBC from NMIBC. Evaluation of metastatic disease is of prime importance, particularly in MIBC, although there are no studies evaluating the use of DWI for this purpose. Whole body DWI is an emerging technique currently being used for other cancers, and Woodhams et al. reported that DWI has the potential in clinical appreciation to examine malignant breast tumours and support the assessment of tumour infiltration [37]. Clinical diagnosis of BC stage is currently done mostly using computed tomography (CT); however, Lista et al. showed that CT could not adequately differentiate the degree of bladder wall infiltration and tumour stage in BC, and its ability to detect invasion to adjacent organs is limited [27]. Takeuchi et al. explained that the accuracy and specificity of DW imaging combined with MRI seem to be higher than those reported previously was due to the enhanced visibility of the structures of the tumour, muscle layer, and thickened submucosa, all of which showed different signal intensity on DW images [12]. In DW-MRI, low signal area with smooth projection is visible in high signal areas, especially in tumour stage 1, which provides a significant advantage of DW-MRI for differentiating NMIBC and MIBC and for revealing the status of muscle layer invasion. Soichiro et al. supported the valuable role of DW-MRI in predicting histologic grade and tumour invasiveness, both of which are key prognostic factors relevant to patient outcome in BC [38].

Limitations of our study should be considered while interpreting our results. First, the quantitative evidence synthesis in our meta-analysis is based on just 11 studies, which threatens the validity of the pooled results. Second, the 11 studies were completed using three different manufacturers of MRIs with field strengths ranging from 1.5 to 3.0 T, which may cause minor heterogeneity. Third, additional high-quality studies may have been omitted because of predefined eligibility criteria or overlooked because studies were published in other languages, lowering the strength of our results.

In conclusion, DW-MRI exhibited outstanding diagnostic performance, with advanced sensitivity and specificity, for imaging of bladder cancers and for differentiating MIBC from NMIBC. We recommend the routine use of DW-MRI for the examination of patients with BC; however, larger sample sizes are still needed to provide a better understanding of the diagnostic performance in relation to long-term survival.

Acknowledgments

We are specifically grateful to the reviewers for their profound comments on this paper.

References

1. Draga ROP, Grimbergen MCM, Vijverberg PLM, van Swol CFP, Jonges TGN, Kummer JA, Bosch JLR. In vivo bladder


