



Published in final edited form as:

*Eur Urol Oncol.* 2020 June ; 3(3): 306–315. doi:10.1016/j.euo.2020.02.007.

## Diagnostic Performance of Vesical Imaging Reporting and Data System for the Prediction of Muscle-invasive Bladder Cancer: A Systematic Review and Meta-analysis

Sungmin Woo<sup>a</sup>, Valeria Panebianco<sup>b</sup>, Yoshifumi Narumi<sup>c</sup>, Francesco Del Giudice<sup>b</sup>, Valdair F. Muglia<sup>d</sup>, Mitsuru Takeuchi<sup>e</sup>, Soleen Ghafoor<sup>a</sup>, Bernard H. Bochner<sup>f</sup>, Alvin C. Goh<sup>f</sup>, Hedvig Hricak<sup>a</sup>, James W.F. Catto<sup>g</sup>, Hebert Alberto Vargas<sup>a,\*</sup>

<sup>a</sup>Department of Radiology, Memorial Sloan Kettering Cancer Center, New York, NY, USA

<sup>b</sup>Department of Radiological, Oncological and Anatomopathological Sciences, “Sapienza” Rome University, Policlinico Umberto I Hospital, Rome, Italy

<sup>c</sup>Departments of Radiology and Health Science, Kyoto Tachibana University, Kyoto, Japan

<sup>d</sup>Imaging Division, Ribeirao Preto Medical School, University of Sao Paulo, Ribeirao Preto, Brazil

<sup>e</sup>Department of Radiology, Radiolonet Tokai, Nagoya, Japan

<sup>f</sup>Urology Service, Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, NY, USA

<sup>g</sup>Academic Urology Unit, University of Sheffield, Sheffield, UK

### Abstract

\*Corresponding author. Department of Radiology, Memorial Sloan-Kettering Cancer Center, 1275 York Ave, Room C-278, New York, NY 10065, USA. Tel: +1-212-639-4017; Fax: +1-212-639-4017. vargasah@mskcc.org (H.A. Vargas).

**Author contributions:** Hebert Alberto Vargas had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

**Study concept and design:** Woo, Panebianco, Narumi, Del Giudice, Muglia, Takeuchi, Ghafoor, Bochner, Goh, Hricak, Catto, Vargas.

**Acquisition of data:** Woo, Vargas.

**Analysis and interpretation of data:** Woo, Panebianco, Narumi, Del Giudice, Muglia, Takeuchi, Ghafoor, Bochner, Goh, Hricak, Catto, Vargas.

**Drafting of the manuscript:** Woo.

**Critical revision of the manuscript for important intellectual content:** Panebianco, Narumi, Del Giudice, Muglia, Takeuchi, Ghafoor, Bochner, Goh, Hricak, Catto, Vargas.

**Statistical analysis:** Woo.

**Obtaining funding:** None.

**Administrative, technical, or material support:** None.

**Supervision:** Vargas.

**Other:** None.

**Financial disclosures:** Hebert Alberto Vargas certifies that all conflicts of interest, including specific financial interests and relationships and affiliations relevant to the subject matter or materials discussed in the manuscript (eg, employment/affiliation, grants or funding, consultancies, honoraria, stock ownership or options, expert testimony, royalties, or patents filed, received, or pending), are the following: Since May 2017, Dr. Hricak has served on the Board of Directors of Ion Beam Applications (IBA), a publicly traded company, and she receives annual compensation for her service. Furthermore, Dr. Hricak is a member of the External Advisory Board of the University of Michigan Comprehensive Cancer Center, the International Advisory Board of the University of Vienna (Austria), and the Scientific Committee of the DKFZ (German Cancer Research Center, Germany); she does not receive financial compensation for any of these roles. Otherwise, the authors do not have any other conflict of interest to disclose.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

**Context:** A noninvasive multiparametric magnetic resonance imaging (MRI)-based scoring system for predicting muscle-invasive bladder cancer (MIBC), the “Vesical Imaging Reporting and Data System” (VI-RADS), was recently developed by an international multidisciplinary panel. Since then, a few studies evaluating the value of VI-RADS for predicting MIBC have been published.

**Objective:** To review the diagnostic performance of VI-RADS for the prediction of MIBC.

**Evidence acquisition:** PubMed and EMBASE databases were searched up to November 10, 2019. We included diagnostic accuracy studies using VI-RADS to predict MIBC using cystectomy or transurethral resection as the reference standard. Methodological quality was evaluated with Quality Assessment of Diagnostic Accuracy Studies-2. Sensitivity and specificity were pooled and plotted using hierarchical summary receiver operating characteristics (HSROC) modeling. Meta-regression analyses were done to explore heterogeneity.

**Evidence synthesis:** Six studies (1770 patients) were included. Pooled sensitivity and specificity were 0.83 (95% confidence interval [CI] 0.70–0.90) and 0.90 (95% CI 0.83–0.95), and the area under the HSROC curve was 0.94 (95% CI 0.91–0.95). Heterogeneity was present among the studies ( $Q = 29.442$ ,  $p < 0.01$ ;  $I^2 = 87.93\%$ , and 90.99% for sensitivity and specificity). Meta-regression analyses showed that the number of patients (>205 vs 205), magnetic field strength (3 vs 1.5 T), T2-weighted image slice thickness (3 vs 4 mm), and VI-RADS cutoff score (3 vs 4) were significant factors affecting heterogeneity ( $p = 0.03$ ).

**Conclusions:** VI-RADS shows good sensitivity and specificity for determining MIBC. Technical factors associated with MRI acquisition and cutoff scores need to be taken into consideration as they may affect performance.

**Patient summary:** A recently established noninvasive magnetic resonance imaging–based scoring system shows good diagnostic performance in detecting muscle-invasive bladder cancer.

## Keywords

Vesical Imaging Reporting and Data System; Magnetic resonance imaging; Bladder cancer; Muscle invasive; Systematic review; Meta-analysis

## 1. Introduction

Bladder cancer is common; there are approximately 550 000 new cases each year worldwide [1]. Determining the extent of bladder wall invasion by tumor is probably the most important component of the initial workup of localized disease, as it directly affects management and prognosis. Tumors demonstrating invasion of the muscularis propria layer (muscle-invasive bladder cancer [MIBC], stage T2) at transurethral resection (TUR) are usually treated with a combination of chemotherapy and radical cystectomy or less frequently radiation. Non-MIBC (NMIBC; stage T1) is usually managed initially with TUR with or without intravesical treatment [2]. Although staging of bladder cancer is established by a combination of clinical, pathological, and imaging means, local staging is primarily based on pathological specimens obtained by TUR, whereas imaging using computed tomography (CT) or magnetic resonance imaging (MRI) has traditionally been more frequently used for the identification of disease outside the bladder (regional lymph nodes, upper tract, or

metastatic disease) [3,4]. Nevertheless, this staging paradigm is suboptimal. A recent meta-analysis found that approximately 10% (95% confidence interval [CI] 6–14%) of T1 bladder cancers were upgraded to MIBC at repeat TUR [5]. In addition, a large-scale population-based study using the US National Cancer Data Base of 18 277 patients with high-grade T1 tumors based on TUR showed that 41% were upstaged at radical cystectomy [6].

The past 2 decades have seen remarkable advances in MRI technology. Implementation of multiparametric MRI (mpMRI), which combines anatomical sequences of T1- (T1WI) and T2-weighted imaging (T2WI) and functional sequences of dynamic contrast-enhanced (DCE) MRI and diffusion-weighted imaging (DWI), has shown incremental value in staging of bladder cancer resulting in high sensitivity and specificity [7,8]. This has led to the development of the Vesical Imaging Reporting and Data System (VI-RADS), which aims to standardize acquisition and reporting of mpMRI for bladder cancer staging with focus on differentiation between MIBC and NMIBC [9]. Since then, several studies evaluating the VI-RADS have been published, but the diagnostic performance of this scoring system is yet to be validated systematically. We conducted a meta-analysis on the diagnostic performance of VI-RADS for predicting MIBC.

## 2. Patients and methods

This meta-analysis was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. A research question was established based on the Patient-Index test- Comparator-Outcome-Study design (PICOS) criteria, which is as follows [10]: What is the diagnostic performance of VI-RADS in predicting MIBC, as compared with pathological results?

### 2.1. Literature search

PubMed and EMBASE databases were systematically searched from inception to November 10, 2019 using a search query established based on keywords of bladder cancer and VI-RADS: (bladder OR urothelial OR papillary OR transitional) AND (cancer OR carcinoma OR tumor OR tumour OR neoplasm) AND (“Vesical Imaging-Reporting And Data System” OR “VI-RADS” OR VIRADS). We additionally screened the bibliographies of initially obtained articles to search for additionally relevant papers.

### 2.2. Study selection

We included studies that met the following PICOS criteria (10): (1) patients (P) had bladder cancer; (2) index test (I) of VI-RADS based on mpMRI was used to determine MIBC, which was the outcome (O) of interest; (3) pathological results based on cystectomy or TUR was used as the reference standard or comparator (C); (4) sufficient data were provided in the study to reconstruct 2 × 2 tables with regard to sensitivity and specificity; and (5) publication type or study design (S) was original article or conference abstract.

The studies were excluded if (1) study population consisted of <10 patients, (2) studies were of other publication types (ie, review articles, editorials, etc.), (3) the study dealt with other topics (eg, MRI used to assess local staging of bladder without VI-RADS, and VI-RADS score for each MRI sequence was provided but an overall VI-RADS score was not tested for

diagnostic performance in determining MIBC), (4) studies had overlapping patient population, and (5) information to reconstruct  $2 \times 2$  tables was insufficient. Authors of the studies were contacted when needed, to request information regarding  $2 \times 2$  tables or other relevant information needed for meta-analytical purposes. With regard to studies with overlapping populations, we selected the study with a greater number of patients and a prospective design (as opposed to a retrospective design).

The literature search and study selection process were performed by two reviewers (initially by S.W. and double checked by H.A.V.).

### 2.3. Data extraction and quality assessment

The following data were extracted from the included studies using a standardized form: origin of study (institution and period of enrollment), size of study population, mean age and range of patients, number of tumors, percentage of patients with MIBC, histological subtypes of tumors, study design (prospective vs retrospective, single- vs multicenter, and whether patients were consecutively enrolled or not), reference standard for determining the status of MIBC, clinical setting of MRI (before vs after TUR), technical parameters of MRI acquisition (including T2WI slice thickness,  $b$  values used for DWI, and temporal resolution of DCE MRI), details regarding MRI interpretation (number and experience of readers, and whether they were blinded or not), and the VI-RADS cutoff value used for determining MIBC on MRI.

We rated the quality of the studies based on the revised Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool [9]. Data extraction and quality assessment were performed by the same two reviewers mentioned above in consensus.

### 2.4. Data synthesis and analysis

Two-by-two contingency tables were constructed from each study based on provided sensitivity and specificity or raw data. For studies that provided diagnostic test accuracy results from multiple readers, the mean value was calculated to represent that study. In addition, if multiple cutoff values were tested within the study, we used the one that provided the highest accuracy or, in other words, the cutoff yielding the least number of false negatives and false positives. Pooled sensitivity and specificity estimates were calculated with hierarchical logistic regression modeling, including bivariate and hierarchical summary receiver operating characteristic (HSROC) modeling, and then graphically presented using HSROC curves with 95% confidence and prediction regions [11,12]. Publication bias was tested by both visual assessment of the Deeks' funnel plot and calculation of  $p$  value using the Deeks' asymmetry test [13].

The presence of heterogeneity was evaluated using several methods: (1) Cochran's  $Q$  test with  $p < 0.05$  signifying heterogeneity; (2) Higgins  $I^2$  test with inconsistency index ( $I^2$ ) = 0–40% (heterogeneity might not be important), 30–60% (moderate heterogeneity), 50–90% (substantial heterogeneity), and 75–100% (considerable heterogeneity) [14]; and (3) threshold effect, a positive correlation between sensitivity and false positive rate among the included studies. To explore potential reasons for heterogeneity, meta-regression analyses were performed for clinically relevant covariates.

We used the “metandi” and “midas” modules in Stata 10.0 (StataCorp LP, College Station, TX, USA) and “mada” package in R software version 3.2.1 (R Foundation for Statistical Computing, Vienna, Austria) for statistical analyses. A *p* value of <0.05 indicated statistical significance except for Deeks’ asymmetry test, where the criterion of <0.1 was used.

### 3. Results

#### 3.1. Literature search

The literature search initially yielded 125 articles. After removing 11 duplicates, screening of the remaining 114 titles and abstracts yielded 20 potentially eligible articles. After full-text review of these articles, 14 studies were excluded for the following reasons: not in field of interest ( $n = 2$ ), overlapping study population ( $n = 1$ ), publication type other than original research or conference abstracts ( $n = 9$ ), and duplicate studies ( $n = 2$ ). Ultimately, six studies (five original articles and one conference abstract) including 1770 patients evaluating the diagnostic performance of VI-RADS for predicting MIBC were included in this meta-analysis, as shown in Figure 1 [15–20].

#### 3.2. Characteristics of included studies

The patient, tumor, and study characteristics are summarized in Table 1. The number of patients ranged from 50 to 297. The prevalence of MIBC ranged from 25% to 50%. Two studies included patients with urothelial cancer only, two included patients predominantly with urothelial cancer but also a few other histological subtypes, and details were not provided in four. Study design was prospective in two studies and retrospective in four. All studies were performed at single centers. All studies used either TUR or a combination of TUR and partial or radical cystectomy as the reference standard. No study solely used cystectomy as the reference standard. In all but one of the above studies, confirmatory repeat TUR was performed for appropriate clinical settings (eg, high- grade NMIBC or insufficient muscle tissue in TUR specimen) to reduce underestimation of MIBC. The time interval between MRI and the reference standard was provided in two studies. MRI was interpreted blinded to the reference standard in all studies.

The MRI characteristics are shown in Table 2. MRI had been performed prior to TUR in all included studies except for one, which consisted of a mixture of pre- and post-TUR patients. Of the studies, 3-T scanners were used in three, 1.5-T scanners in one, and either 1.5- or 3-T scanners in one, and details were not provided in one. All but one study in which details were not available generally abided by the recommendations provided in the VI-RADS guidelines for image acquisition [9]. For example, T2-weighted images were obtained with a slice thickness of 3–4 mm and diffusion-weighted images were acquired with high *b* values of 1000 s/mm<sup>2</sup> or higher. Although protocols for DCE MRI varied among studies, temporal resolution was sufficient for the depiction of early enhancement of the inner layer followed by tumor enhancement required for VI-RADS evaluation [21]. In all studies, MRI was interpreted blinded to the pathological findings. A cutoff VI-RADS score of 3 or 4 was used to determine MIBC.

### 3.3. Quality assessment

The quality of the studies were moderate to good, with all studies satisfying at least four out of the seven QUADAS-2 domains (Fig. 2). Regarding the patient selection domain, none of the studies had any risk of bias. In the index test domain, there was a high risk of bias in two studies [16,18] as the threshold (VI-RADS cutoff score for determining MIBC) was not prespecified, but determined based on receiver operating characteristics curves and partition analysis, while in another two [17,19], it was unclear whether they were prespecified or not. Regarding the reference standard domain, two studies [18,20] had an unclear risk of bias as the reference standard included TUR, but it was not clarified whether a confirmatory or secondary repeat TUR was performed in clinically relevant situations (eg, no muscle tissue included in the TUR specimen). Regarding the flow and timing domain, three studies had a high risk of bias as different reference standards (cystectomy and TUR) were applied to patients within the study [15,16,20]. Three were considered to have an unclear risk of bias as the interval between mpMRI acquisition used for VI-RADS scoring and reference standard was not provided [17–19]. There was low concern for applicability in all QUADAS-2 domains for all studies.

### 3.4. Diagnostic accuracy of VI-RADS for the prediction of MIBC

The pooled sensitivity and specificity for predicting MIBC were 0.83 (95% CI 0.70–0.90) and 0.90 (95% CI 0.83–0.95). The area under the HSROC curve was 0.94 (95% CI 0.91–0.95; Fig. 3). Based on the Cochran's  $Q$  test, heterogeneity was present ( $Q = 29.442$ ,  $p < 0.01$ ). Higgins  $I^2$  statistics indicated that there was considerable heterogeneity regarding both sensitivity ( $I^2 = 87.93\%$ ) and specificity ( $I^2 = 90.99\%$ ). There was no threshold effect based on visual assessment of the coupled forest plot of sensitivity and specificity (Fig. 4), and the calculated correlation coefficient between sensitivity and false positive rate of 0.110 (95% CI –0.770 to 0.846). Publication bias was not suggested based on the Deeks' funnel plot, with a  $p$  value of 0.52 for slope coefficient (Fig. 5).

### 3.5. Heterogeneity exploration

Meta-regression analyses showed that the number of patients, magnetic field strength, T2WI slice thickness, and cutoff value used for determining MIBC were significant covariates affecting the heterogeneity ( $p = 0.03$ ). Among them, only sensitivity stratified to the number of included patients showed clinical and statistically significant difference: 0.90 (95% CI 0.86–0.95) in studies with  $>205$  patients versus 0.67 (0.55–0.78) in studies with  $<205$  patients (Table 3). The following parameters showed substantial differences in pooled estimates with minimal overlap in their 95% CIs: (1) magnetic field strength (3 vs 1.5 T), sensitivity of 0.90 (95% CI 0.86–0.94) versus 0.77 (95% CI 0.65–0.88); (2) T2WI slice thickness (3 vs 4 mm), sensitivity of 0.90 (95% CI 0.86–0.95) versus 0.84 (95% CI 0.77–0.90) and specificity of 0.85 (95% CI 0.78–0.93) versus 0.96 (95% CI 0.93–1.00); and (3) VI-RADS score cutoff values (4 vs 3), sensitivity of 0.72 (95% CI 0.61–0.83) versus 0.92 (95% CI 0.86–0.97) and specificity of 0.96 (95% CI 0.91–1.00) versus 0.84 (95% CI 0.65–1.00). Other variables were not significant factors related to heterogeneity ( $p = 0.09–0.76$ ).

### 3.6. Discussion

In this meta-analysis, we assessed the diagnostic performance of VI-RADS for predicting MIBC. This is the first study to systematically review and meta-analyze the currently available evidence in the literature dealing with this new scoring system for bladder MRI. We found that the pooled sensitivity and specificity for predicting MIBC were 0.83 and 0.90, respectively. This diagnostic performance of VI-RADS is similar to the diagnostic performance of bladder MRI in determining MIBC prior to the introduction of VI-RADS based on a previous meta-analysis of 24 studies, in which the pooled sensitivity and specificity were 0.92 (95% CI 0.88–0.95) and 0.87 (95% CI 0.78–0.93) [7]. VI-RADS offers the added value of establishing a standardized approach to both acquisition and reporting of mpMRI for bladder cancer. With regard to acquisition, all studies but one conference abstract (which did not provide sufficient details) abided by the acquisition guidelines (eg, magnetic field strength, slice thickness on T2WI, and temporal resolution of DCE MRI). In addition, in most of the included studies, there was substantial inter-reader agreement, with kappa values ranging from 0.81 to 0.92 [15–17,20] and intraclass correlation coefficient of 0.85 [19]. In addition, standardized interpretation and reporting, which are based on accumulated evidence in the literature, could potentially increase the performance of less experienced readers, as in the case with Prostate Imaging Reporting and Data System compared with Likert scales [22]. This high degree of agreement along with good diagnostic performance could potentially be the cornerstone of bladder MRI and VI-RADS to be implemented in clinical practice and could, in turn, add incremental value to the currently suboptimal staging paradigm, where changes from NMIBC on TUR to MIBC on radical cystectomy can be seen up to 41% of patients [6].

Considerable degree of heterogeneity was seen among the studies, and meta-regression analyses provided some insight regarding potential causes for this heterogeneity. Among several potential factors, the number of patients included in each study was the only statistically significant factor. Specifically, studies that included a greater number of patients demonstrated significantly higher sensitivity: 0.90 (95% CI 0.86–0.95) in studies with >205 patients versus 0.67 (0.55–0.78) in studies with <205 patients. Part of this may be related to experience in VI-RADS or interpretation of bladder mpMRI in general. Capability to include a larger number of patients possibly indicates a higher-volume center with potentially greater exposure of the affiliated radiologists to bladder MRI. In addition, level of experience is shown to affect the diagnostic performance of bladder MRI in differentiating MIBC versus NMIBC, with radiologists with greater expertise showing higher accuracy than those with lesser experience [23].

A few technical aspects of MRI acquisition were also associated with heterogeneity. Although there was minimal overlap in the 95% CIs, (1) studies using only 3-T scanners demonstrated substantially higher sensitivity than those using 1.5-T scanners (0.90 vs 0.77) and (2) studies using 3-mm slice thickness demonstrate substantially higher sensitivity (0.90 vs 0.84) and lower specificity (0.85 vs 0.96) than those using 4-mm slice thickness. The use of 3-T scanners offers substantially higher signal-to-noise and contrast-to-noise ratios than 1.5-T scanners, improving differentiation of the layers of the bladder wall, which is essential for assigning VI-RADS scores [24]. In addition, the use of thinner slice thickness

contributes to reducing partial volume averaging [25]. Such technical factors have consistently been shown in previous meta-analyses to affect significantly the diagnostic performance of local staging of tumors in the pelvis, including bladder cancer (prior to using VI-RADS), cervical cancer, and rectal cancer [7,26,27].

Another important factor affecting heterogeneity was the VI-RADS cutoff score. Albeit minimal overlap in 95% CI, sensitivity was substantially higher (0.96 vs 0.72) and specificity was lower (0.84 vs 0.92) when using score of 22653 as opposed to 4 for determining MIBC. This rather intuitive and expected heterogeneity stemming from the fundamental nature of all diagnostic tests could be applied differentially according to the clinical scenario. For example, a VI-RADS cutoff score of 3 (rather than 4) may be used in situations where greater sensitivity for the detection of MIBC is favored, such as in patients with known history of bladder cancer who are at a higher risk of developing MIBC (eg, recurrent, multiple, larger [ $>3$  cm], higher T stage [pT1 vs pTa], concomitant carcinoma in situ, and higher grade [G3 vs G1 or G2]) [28], or in patients with upper tract urothelial cancers under surveillance after nephroureterectomy who are at an increased risk of developing MIBC (eg, higher stage [ pT3 vs Ta or T1] and distal location [distal ureter vs renal pelvis]) [29].

There are some limitations in our meta-analysis. First, the number of included studies was small ( $n = 6$ ) as VI-RADS has only recently been introduced, and there is some heterogeneity among the included studies. Nevertheless, it included a relatively large number of patients ( $n = 1770$ ) and therefore represents the available evidence as of now regarding its diagnostic performance to predict MIBC. Second, all studies were performed at single centers and most were retrospective by design. In order to test the validity of VI-RADS, especially for assessing its potential for greater reproducibility, prospective multicenter trials are warranted. Third, only TUR, or either TUR or cystectomy was used as the reference standard. Unlike cystectomy, TUR has the potential to undersample the muscle layer, resulting in less accurate determination of MIBC. Nevertheless, all but one of the studies, which utilized TUR as the reference standard, performed secondary TUR in appropriate clinical settings (eg, high-grade MIBC or repeat TUR if no muscle tissue was present in TUR sample). Still, additional confirmatory TUR holds the possibility of understaging, and studies using only cystectomy as the reference standard or comparing TUR, cystectomy, and MRI findings will be of added value in determining the diagnostic performance of VI-RADS [30]. Fourth, all but one study was based on patients in the pre-TUR setting only, which showed sensitivity and specificity of 0.81 (0.70–0.92) and 0.87 (0.82–0.93), respectively; therefore, we cannot reach a firm conclusion as of whether the diagnostic performance of VI-RADS in this meta-analysis will be applicable to patients who undergo MRI after TUR. It has been shown that postoperative inflammation and fibrosis in the bladder wall after TUR can mimic tumor presence or obscure small recurrent tumors. For example, the diagnostic performance of using solely T2WI for differentiating residual tumor versus post-TUR inflammation is approximately 50% [31]. Although there is ample evidence in the literature showing that the use of mpMRI, incorporating DWI and DCE MRI, can detect recurrent tumors with higher diagnostic performance than T2WI [32–34], it is not well known whether this extrapolates to good performance in determining MIBC, specifically in terms of the application of VI-RADS. Currently, VI-RADS does not



specifically address this question, and future studies are warranted to evaluate VI-RADS in various clinical settings, including MRI performed after TUR, when considering candidates for trimodal bladder-sparing therapy, or when patients are under surveillance for tumor recurrence. Finally, it is important to emphasize that the purpose of VI-RADS is to detect MIBC; it cannot evaluate for the presence of carcinoma in situ, which is considered an important prognostic factor in currently available prediction models.

#### 4. Conclusions

VI-RADS shows good performance for the prediction of MIBC, with pooled sensitivity of 0.83 and specificity of 0.90. Technical factors such as magnetic field strength and slice thickness, and cutoff values may be associated with heterogeneity. Application of these results in post-TUR patients is limited due to the paucity of studies done in this setting.

#### Acknowledgments

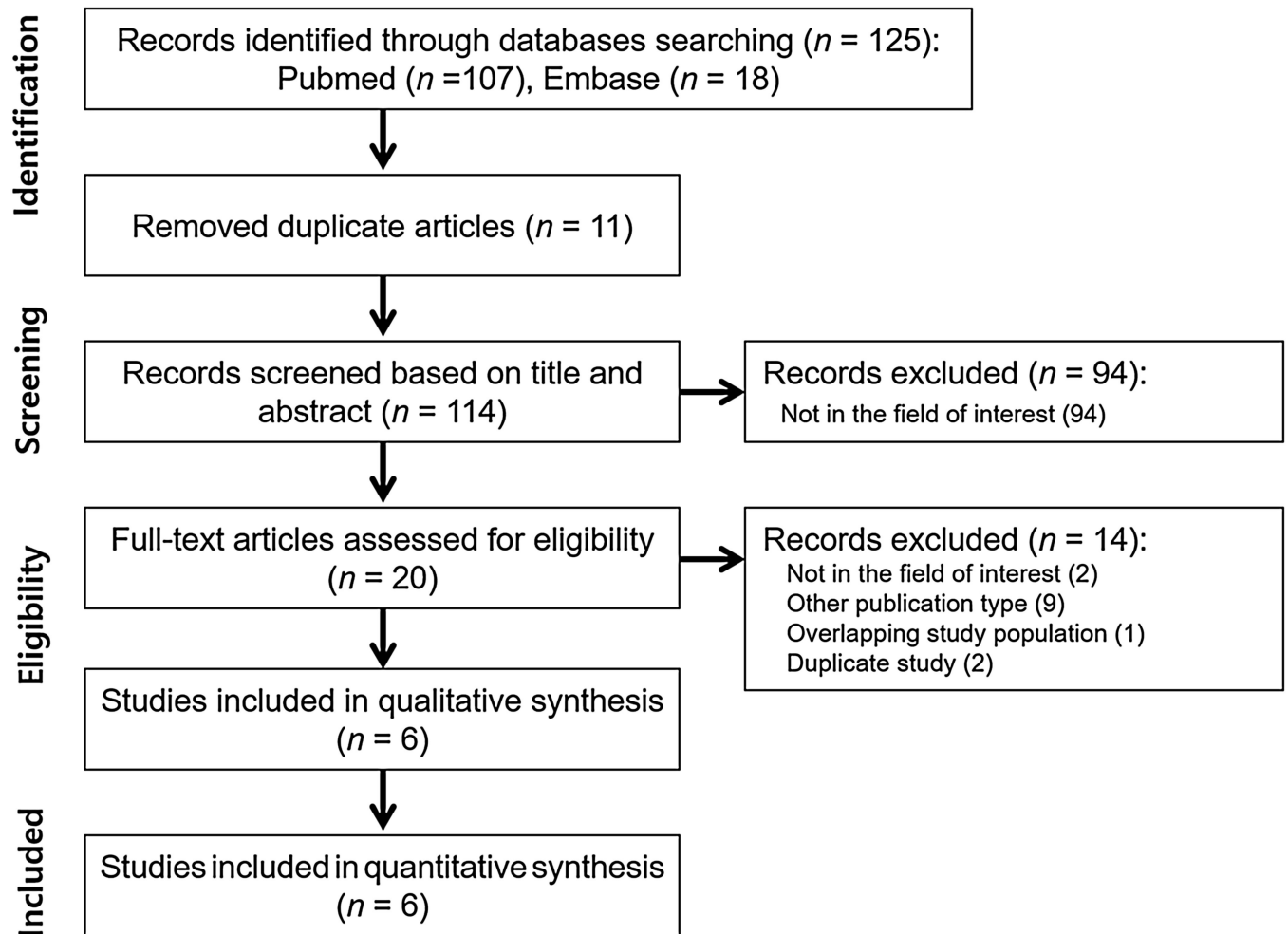
**Funding/Support and role of the sponsor:** This research was funded in part through the NIH/NCI Cancer Center Support Grant P30 CA008748. However, all specific roles were performed only by authors, and NIH/NCI had no direct role or support in those roles.

#### References

- [1]. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018;68:394–424. [PubMed: 30207593]
- [2]. Josephson D, Pasin E, Stein JP. Superficial bladder cancer: part 2. Management. *Expert Rev Anticancer Ther* 2007;7:567–81. [PubMed: 17428176]
- [3]. Chang SS, Boorjian SA, Chou R, et al. Diagnosis and treatment of non-muscle invasive bladder cancer: AUA/SUO guideline. *J Urol* 2016;196:1021–9. [PubMed: 27317986]
- [4]. Babjuk M, Burger M, Comperat EM, et al. European Association of Urology guidelines on non-muscle-invasive bladder cancer (TaT1 and carcinoma in situ)—2019 update. *Eur Urol* 2019;76:639–57. [PubMed: 31443960]
- [5]. Naselli A, Hurler R, Paparella S, et al. Role of restaging transurethral resection for T1 non-muscle invasive bladder cancer: a systematic review and meta-analysis. *Eur Urol Focus* 2018;4:558–67. [PubMed: 28753839]
- [6]. Matulewicz RS, Franey BT, Oberlin DT, Meeks JJ. High-risk of adverse pathologic features in patients with clinical T1 high-grade bladder cancer undergoing radical cystectomy. *J Natl Compr Canc Netw* 2016;14:1403–11. [PubMed: 27799511]
- [7]. Woo S, Suh CH, Kim SY, Cho JY, Kim SH. Diagnostic performance of MRI for prediction of muscle-invasiveness of bladder cancer: a systematic review and meta-analysis. *Eur J Radiol* 2017;95:46–55. [PubMed: 28987698]
- [8]. Vargas HA, Akin O, Schoder H, et al. Prospective evaluation of MRI, (1)(1)C-acetate PET/CT and contrast-enhanced CT for staging of bladder cancer. *Eur J Radiol* 2012;81:4131–7. [PubMed: 22858427]
- [9]. Panebianco V, Narumi Y, Altun E, et al. Multiparametric magnetic resonance imaging for bladder cancer: development of VI-RADS (Vesical Imaging-Reporting and Data System). *Eur Urol* 2018;74:294–306. [PubMed: 29755006]
- [10]. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol* 2009;62:e1–34. [PubMed: 19631507]
- [11]. Suh CH, Park SH. Successful publication of systematic review and meta-analysis of studies evaluating diagnostic test accuracy. *Korean J Radiol* 2016;17:5–6. [PubMed: 26798211]

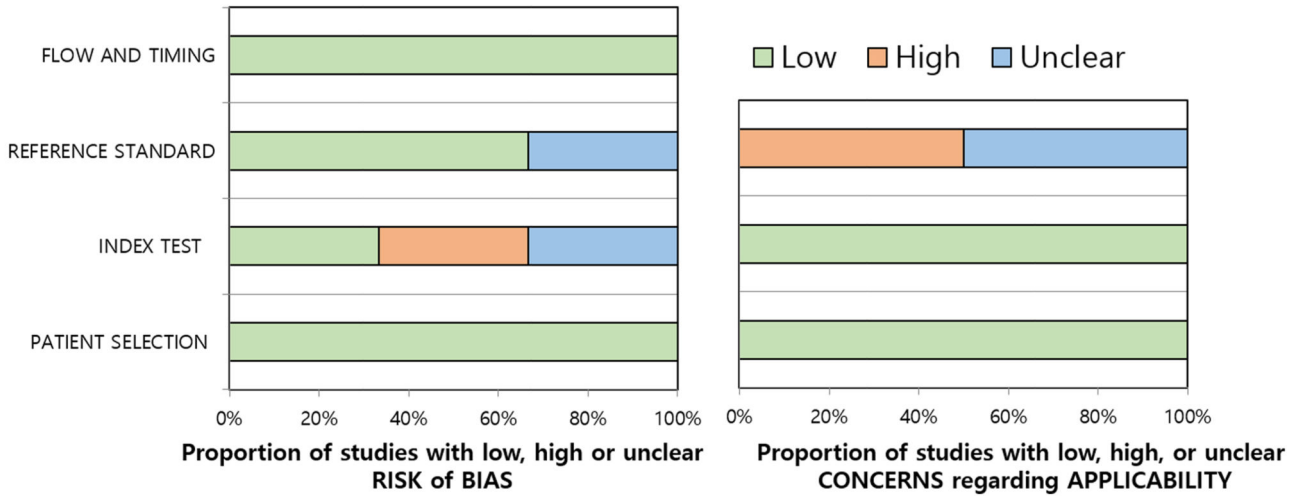
- [12]. Lee J, Kim KW, Choi SH, Huh J, Park SH. Systematic review and meta-analysis of studies evaluating diagnostic test accuracy: a practical review for clinical researchers—part II. Statistical methods of meta-analysis. *Korean J Radiol* 2015;16:1188–96. [PubMed: 26576107]
- [13]. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol* 2005;58:882–93. [PubMed: 16085191]
- [14]. Higgins J, Green S. *Cochrane handbook for systematic reviews of interventions version 5.1.0*. The Cochrane Collaboration. [http://handbook.cochrane.org/chapter\\_9/9\\_5\\_2\\_identifying\\_and\\_measuring\\_heterogeneity.htm](http://handbook.cochrane.org/chapter_9/9_5_2_identifying_and_measuring_heterogeneity.htm)
- [15]. Del Giudice F, Barchetti G, De Berardinis E, et al. Prospective assessment of Vesical Imaging Reporting and Data System (VI-RADS) and its clinical impact on the management of high-risk non-muscle-invasive bladder cancer patients candidate for repeated transurethral resection. *Eur Urol* 2020;77:101–9. [PubMed: 31699526]
- [16]. Kim SH. Validation of vesical imaging reporting and data system for assessing muscle invasion in bladder tumor. *Abdom Radiol (NY)* 2020;45:491–8. [PubMed: 31422440]
- [17]. Makboul M, Farghaly S, Abdelkawi IF. Multiparametric MRI in differentiation between muscle invasive and non-muscle invasive urinary bladder cancer with vesical imaging reporting and data system (VI-RADS) application. *Br J Radiol* 2019;92:20190401.
- [18]. Sakamoto K, Ito M, Nakanishi Y, et al. Prediction of muscle invasive bladder cancer using the Vesical Imaging-Reporting and Data System and apparent diffusion coefficient values (VI-RADS/ADC). *Eur Urol Suppl* 2019;18:e242–3.
- [19]. Ueno Y, Takeuchi M, Tamada T, et al. Diagnostic accuracy and interobserver agreement for the Vesical Imaging-Reporting and Data System for muscle-invasive bladder cancer: a multireader validation study. *Eur Urol* 2019;76:54–6. [PubMed: 30922688]
- [20]. Wang H, Luo C, Zhang F, et al. Multiparametric MRI for bladder cancer: validation of VI-RADS for the detection of detrusor muscle invasion. *Radiology* 2019;291:668–74. [PubMed: 31012814]
- [21]. Hayashi N, Tochigi H, Shiraishi T, Takeda K, Kawamura J. A new staging criterion for bladder carcinoma using gadolinium-enhanced magnetic resonance imaging with an endorectal surface coil: a comparison with ultrasonography. *BJU Int* 2000;85:32–6.
- [22]. Catala V, Barcina MJM, Mayordomo O, et al. Characterization of prostate lesions as benign or malignant by multiparametric 3 T MR imaging: comparison of Likert score to the Prostate Imaging Reporting and Data System 2 version. *European Congress of Radiology*; 2016.
- [23]. Wu LM, Chen XX, Xu JR, et al. Clinical value of T2-weighted imaging combined with diffusion-weighted imaging in preoperative T staging of urinary bladder cancer: a large-scale, multiobserver prospective study on 3.0-T MRI. *Acad Radiol* 2013;20:939–46. [PubMed: 23746384]
- [24]. Chang KJ, Kamel IR, Macura KJ, Bluemke DA. 3.0-T MR imaging of the abdomen: comparison with 1.5 T. *Radiographics* 2008;28:1983–98. [PubMed: 19001653]
- [25]. Zand KR, Reinhold C, Haider MA, Nakai A, Rohoman L, Maheshwari S. Artifacts and pitfalls in MR imaging of the pelvis. *J Magn Reson Imaging* 2007;26:480–97. [PubMed: 17623875]
- [26]. Woo S, Suh CH, Kim SY, Cho JY, Kim SH. Magnetic resonance imaging for detection of parametrial invasion in cervical cancer: an updated systematic review and meta-analysis of the literature between 2012 and 2016. *Eur Radiol* 2018;28:530–41. [PubMed: 28726120]
- [27]. Kim TH, Woo S, Han S, Suh CH, Vargas HA. The diagnostic performance of MRI for detection of extramural venous invasion in colorectal cancer: a systematic review and meta-analysis of the literature. *AJR Am J Roentgenol* 2019;213:575–85. [PubMed: 31063424]
- [28]. Sylvester RJ, van der Meijden AP, Oosterlinck W, et al. Predicting recurrence and progression in individual patients with stage Ta T1 bladder cancer using EORTC risk tables: a combined analysis of 2596 patients from seven EORTC trials. *Eur Urol* 2006;49:466–75; discussion 475–7. [PubMed: 16442208]
- [29]. Kim KH, You D, Jeong IG, Hong JH, Ahn H, Kim CS. Muscle-invasive bladder cancer developing after nephroureterectomy for upper urinary tract urothelial carcinoma. *Urol Oncol* 2013;31:1643–9. [PubMed: 22591745]

- [30]. Ark JT, Keegan KA, Barocas DA, et al. Incidence and predictors of understaging in patients with clinical T1 urothelial carcinoma undergoing radical cystectomy. *BJU Int* 2014;113:894–9. [PubMed: 24053444]
- [31]. Nakamura Y, Yoshida S, Tanaka H, et al. Potential utility of diffusion-weighted magnetic resonance imaging in diagnosis of residual bladder cancer before second transurethral resection. *Urol Int* 2017;98:298–303. [PubMed: 28245479]
- [32]. Rosenkrantz AB, Ego-Osuala IO, Khalef V, Deng FM, Taneja SS, Huang WC. Investigation of multisequence magnetic resonance imaging for detection of recurrent tumor after transurethral resection for bladder cancer. *J Comput Assist Tomogr* 2016;40:201–5. [PubMed: 26760195]
- [33]. El-Assmy A, Abou-El-Ghar ME, Refaie HF, Mosbah A, El-Diasty T. Diffusion-weighted magnetic resonance imaging in follow-up of superficial urinary bladder carcinoma after transurethral resection: initial experience. *BJU Int* 2012;110:E622–7. [PubMed: 22757606]
- [34]. Wang HJ, Pui MH, Guo Y, Yang D, Pan BT, Zhou XH. Diffusion-weighted MRI in bladder carcinoma: the differentiation between tumor recurrence and benign changes after resection. *Abdom Imaging* 2014;39:135–41. [PubMed: 24072383]

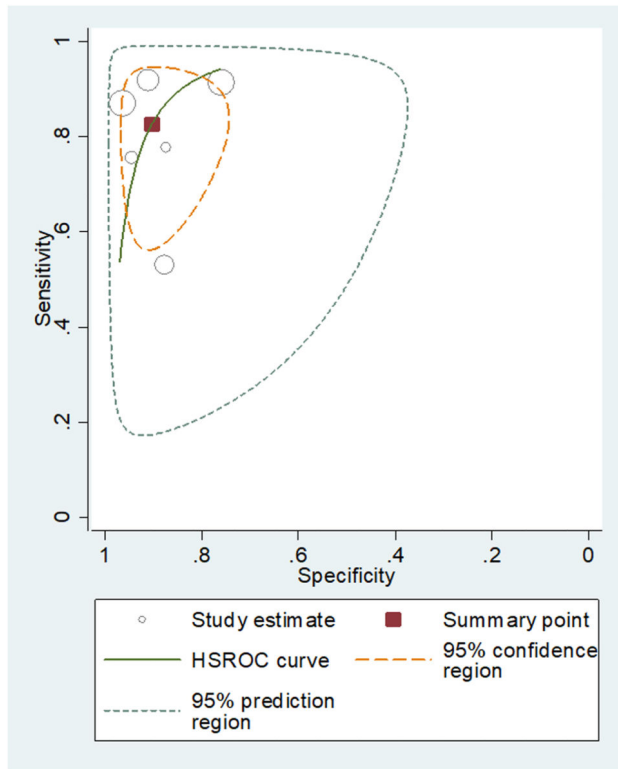


**Fig. 1** -.  
Study selection process.

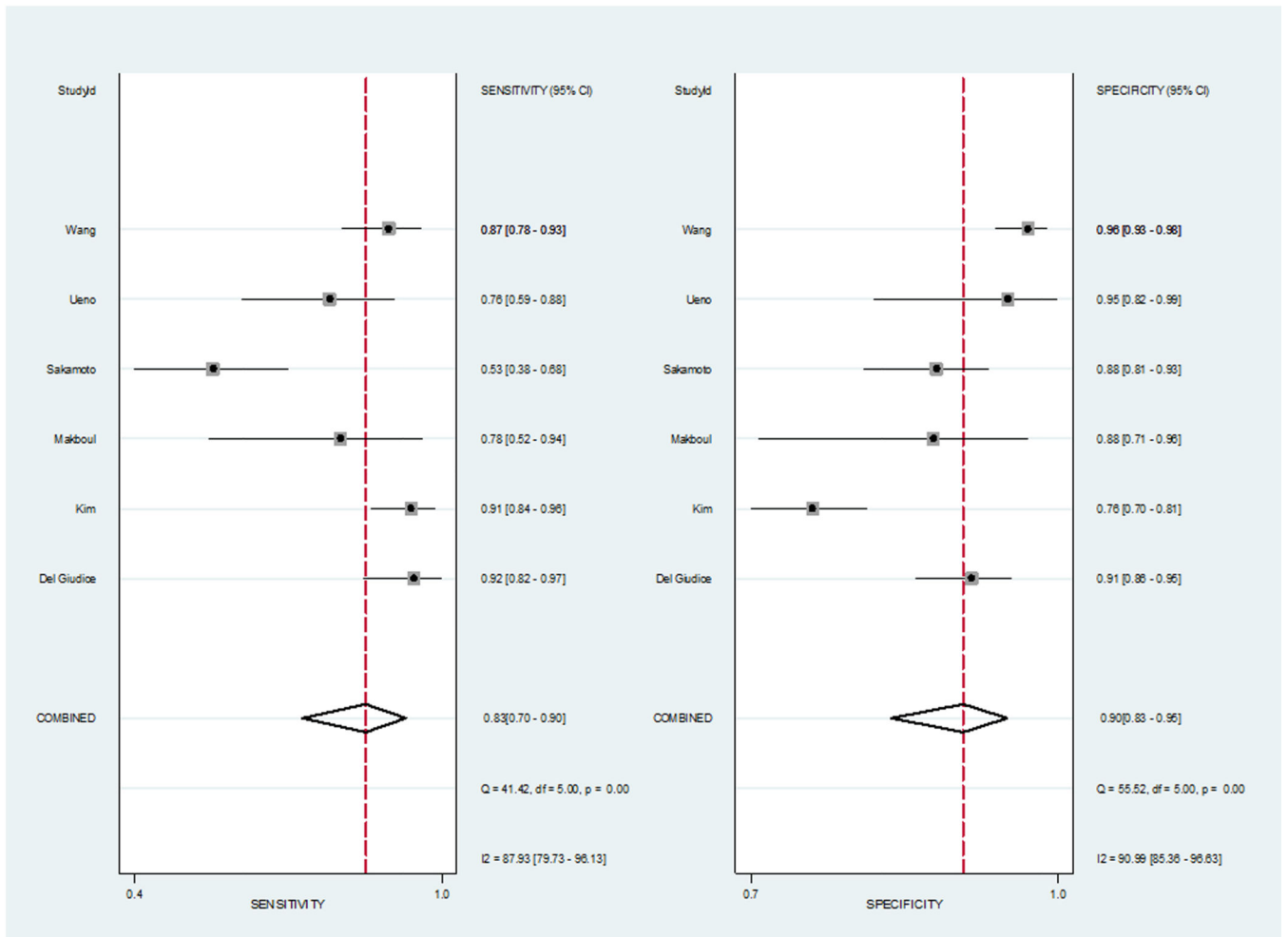
**QUADAS-2 Domain**



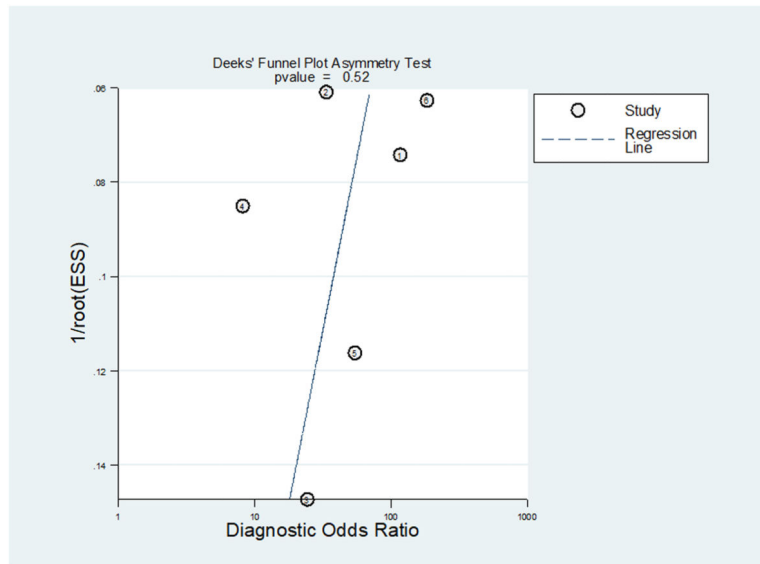
**Fig. 2 –.** Grouped bar charts show (A) risk of bias and (B) concerns for applicability of six studies included in meta-analysis using QUADAS-2. QUADAS-2 = Quality Assessment of Diagnostic Accuracy Studies-2.



**Fig. 3 –.** Hierarchical summary receiver operating characteristic curve showing diagnostic performance of studies using VI-RADS for the prediction of muscle-invasive bladder cancer. HSROC = hierarchical summary receiver operating characteristics; VI-RADS = Vesical Imaging Reporting and Data System.



**Fig. 4 –.** Coupled forest plots of sensitivity and specificity. Numbers are pooled estimates with 95% confidence intervals (CIs) in parentheses and heterogeneity statistics are shown at bottom right. Horizontal lines indicate 95% CIs. df = degree of freedom.



**Fig. 5.**  
– Deeks' funnel plot. A  $p$  value of 0.52 indicates absence of publication bias. ESS = effective sample size.



Table 1 –

Patient, tumor, and study characteristics

Author (year) [reference no.]	Patient characteristics				Tumor characteristics				Study design			Reference standard	
	Period	Institution	No.	Age (yr), median (range)	No.	Pathological T stage		Histological subtype (no.)	Prospective	Multicenter	Consecutive enrollment	Method	MRI reference standard interval
						Range	T2 (%)						
Del Giudice et al (2020) [15]	Sep 2017–May 2019	Sapienza University of Rome	231	65.5 (47–79)	103	Ta-T4	27	All urothelial	Yes	No	Yes	TUR <sup>a</sup> , RC	<6 wk
Kim (2020) [16]	Jan 2015–Mar 2019	Kyungpook National University Hospital	297	65.5 <sup>b</sup> (58–75)	339	Tis-T4	34	Urothelial (320), urothelial with adenocarcinoma differentiation (14), undifferentiated (5)	No	No	Yes	TUR <sup>a</sup> , PC, RC	NR
Makboul et al (2019) [17]	NR	Assiut University Hospital	50	57.2 <sup>b</sup> (NR)	50	T1-T4	36	NR	Yes	No	Yes	TUR <sup>a</sup>	NR
Sakamoto et al (2019) [18]	2013–2018	Tokyo Metropolitan Cancer and Infectious Diseases Center Komagome Hospital	178	NR	178	NR	26	NR	No	No	Yes	TUR	NR
Ueno et al (2019) [19]	2010–2018	Kobe University Graduate School of Medicine	74	NR	74	Tis-T2	50	Urothelial (71), small cell (1), squamous cell (2)	No	No	Yes	TUR <sup>a</sup>	NR
Wang et al (2019) [20]	Nov 2011–Aug 2018	First Affiliated Hospital of Sun Yat-Sen University	340	64 (27–87)	340	Ta-T4	25	All urothelial	No	No	Yes	TUR, PC, RC	<2 wk

MIBC = muscle-invasive bladder cancer; MRI = magnetic resonance imaging; NR = not reported; PC = partial cystectomy; RC = radical cystectomy; TUR = transurethral resection

<sup>a</sup>Confirmatory repeat TUR when clinically indicated (eg, high-grade non-MIBC, no muscle tissue in TUR specimen).

<sup>b</sup>Mean.

Table 2 –

## MRI characteristics

Author (year) [reference no.]	Setting	Technical parameters				Reader characteristics			
		Magnet strength (Tesla)	T2WI slice thickness (mm)	DWI <i>b</i> values (s/mm <sup>2</sup> )	DCE MRI temporal resolution (s)	Number	Consensus	Reader experience (yr)	Blinded
Del Giudice et al (2020) [15]	Pre-TUR	3	3–4	0, 800, 1000, 2000	Every 5 s	2	Consensus	>10, 5	Yes
Kim (2020) [16]	Pre-TUR	3	3	0, 1000	Every 30 s for 4–6 acquisitions	2	Consensus	12, 6	Yes
Makboul et al (2019) [17]	Pre-TUR	1.5	3	0, 400, 800, 1000	At 20, 70, 180 s	2	Consensus	NR	Yes
Sakamoto et al (2019) [18]	Pre-TUR	NR	NR	NR	NR	NR	NR	NR	Yes
Ueno et al (2019) [19]	Pre-TUR	1.5, 3	4	0, 1000	At 40, 80, 120, 160, 200 s	5	Independent	NR	Yes <sup>a</sup>
Wang et al (2019) [20]	Pre- and post-TUR	3	4	0, 1000	5 acquisitions between 20 and 131 s	2	Consensus	32, 8	Yes

DCE = dynamic contrast enhanced; DWI = diffusion-weighted imaging; MRI = magnetic resonance imaging; NR = not reported; TUR = transurethral resection; T2WI = T2-weighted imaging.

<sup>a</sup>Blinded to pathological tumor stage but informed on the location of tumor.

**Table 3 –**

## Meta-regression analysis

Covariate		No. of studies	Sensitivity (95% CI)	Specificity (95% CI)	<i>p</i> value
No. of patients	>205	3	0.90 (0.86–0.95)	0.90 (0.83–0.97)	0.02
	205	3	0.67 (0.55–0.78)	0.90 (0.82–0.99)	
Proportion of MIBC (%)	>30	3	0.84 (0.71–0.97)	0.85 (0.76–0.95)	0.42
	30	3	0.81 (0.67–0.95)	0.93 (0.88–0.97)	
Study design	Prospective	2	0.88 (0.74–1.00)	0.90 (0.81–1.00)	0.76
	Retrospective	4	0.80 (0.68–0.92)	0.90 (0.84–0.97)	
Clinical setting for MRI	Pre-TUR	5	0.81 (0.70–0.92)	0.87 (0.82–0.93)	0.09
	Pre- and post-TUR	1	0.87 (0.82 – 0.93)	0.97 (0.93 – 1.00)	
Magnetic strength	3 T only	3	0.90 (0.86–0.94)	0.90 (0.83–0.98)	<0.01
	1.5 T used	2	0.77 (0.65–0.88)	0.92 (0.82–1.00)	
T2WI slice thickness (mm)	3	3	0.90 (0.86–0.95)	0.85 (0.78–0.93)	<0.01
	4	2	0.84 (0.77–0.90)	0.96 (0.93–1.00)	
VI-RADS cutoff score	4	5	0.72 (0.61 –0.83)	0.96 (0.91–1.00)	0.03
	3	4	0.92 (0.86 –0.97)	0.84 (0.65–1.00)	

CI = confidence interval; MIBC = muscle-invasive bladder cancer; MRI = magnetic resonance imaging; TUR = transurethral resection; T2WI = T2-weighted imaging; VI-RADS = Vesical Imaging Reporting and Data System.