AJNR

The Diagnostic Value of Diffusion-Weighted Imaging in Differentiating Metastatic Lymph Nodes of Head and Neck Squamous Cell Carcinoma: A Systematic Review and Meta-Analysis

C.H. Suh, Y.J. Choi, J.H. Baek and J.H. Lee

The Diagnostic Value of Diffusion-Weighted Imaging in Differentiating Metastatic Lymph Nodes of Head and Neck Squamous Cell Carcinoma: A Systematic Review and Meta-Analysis

[ID]C.H. Suh, [ID]Y.J. Choi, [ID]J.H. Baek, and [ID]J.H. Lee

ABSTRACT

BACKGROUND: Accurate lymph node staging is crucial for proper treatment planning for metastasis in patients with head and neck squamous cell carcinoma.

PURPOSE: Our aim was to evaluate the diagnostic performance of DWI for differentiating metastatic cervical lymph nodes from benign cervical lymph nodes in patients with head and neck squamous cell carcinoma and to identify optimal cutoff values for ADC.

DATA SOURCES: A computerized literature search was performed to identify relevant original articles in Ovid MEDLINE and EMBASE.

STUDY SELECTION: Studies evaluating the diagnostic performance of DWI for differentiating metastatic cervical lymph nodes from benign cervical lymph nodes were selected.

DATA ANALYSIS: Diagnostic meta-analysis was conducted with a bivariate random-effects model, and a hierarchical summary receiver operating characteristic curve was obtained. Meta-regression was also performed.

DATA SYNTHESIS: Nine studies with 337 patients were included. In all studies, ADC values derived from metastatic lymph nodes were significantly lower than ADC values derived from benign lymph nodes. The median ADC cutoff value was 0.965×10^{-3} mm^2/s. The pooled sensitivity and specificity for the diagnostic performance of DWI in differentiating metastatic lymph nodes from benign lymph nodes were 90% (95% CI, 84%–94%) and 88% (95% CI, 80%–93%), respectively. In the meta-regression, sensitivity was significantly higher in the studies using a 3-mm slice thickness (93% [95% CI, 88%–98%]) than in studies using a slice thickness of >3 mm (86% [95% CI, 77%–95%], $P < .01$).

LIMITATIONS: A small number of studies were included in our meta-analysis.

CONCLUSIONS: DWI demonstrated high diagnostic performance for differentiating metastatic lymph nodes from benign lymph nodes in patients with head and neck squamous cell carcinoma, and the median ADC cutoff value was 0.965×10^{-3} mm^2/s. A 3-mm DWI slice thickness can provide a slight improvement in sensitivity.

ABBREVIATIONS: HNSCC = head and neck squamous cell carcinoma; LN = lymph node; NCCN = National Comprehensive Cancer Network; QUADAS-2 = Quality Assessment of Diagnostic Accuracy Studies-2

Lymph node (LN) metastasis is an adverse prognostic factor in patients with head and neck squamous cell carcinoma (HNSCC), and accurate LN staging is crucial for proper treatment planning. The National Comprehensive Cancer Network (NCCN) guidelines recommend CT and/or MR imaging with contrast for the initial work-up in patients with HNSCC.[1] CT and MR imaging are useful for determining morphologic criteria, including shape, size, internal architecture, extracapsular extension, and vascular features associated with LN metastasis; however, diagnostic performance is limited, especially in normal-sized non-necrotic LNs.[2,3]

During the past decade, diffusion-weighted imaging has been used for differentiating metastatic cervical LNs from benign cervical LNs in patients with HNSCC. Some studies have reported a high diagnostic performance for DWI,[4-12] whereas other studies have demonstrated disappointing results.[13,14] In addition, various cutoff values have been proposed for the apparent diffusion coefficient.

We considered it timely to review the DWI protocols, parameters, and reported diagnostic performances because there are no

published systematic reviews or meta-analyses on the topic, to our knowledge. Therefore, we aimed to evaluate the diagnostic performance of DWI for differentiating metastatic cervical LNs from benign cervical LNs in patients with HNSCC and to identify the optimal ADC cutoff value.

MATERIALS AND METHODS

This study adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines.[15]

Search Strategy

A preliminary literature search demonstrated that various terminology was used to indicate cervical LN metastasis, including "cervical nodal metastasis," "malignant cervical lymph nodes," "metastatic cervical lymph node," and "cervical lymphadenopathy." These synonyms for cervical LN metastasis were used in the search terms for Ovid MEDLINE and EMBASE: ((cervical lymph node metastasis) OR (cervical nodal metastasis) OR (malignant cervical lymph nodes) OR (metastatic cervical lymph node) OR (cervical lymphadenopathy)) AND ((diffusion-weighted) OR (DWI) OR (apparent diffusion coefficient) OR (ADC)). The literature search was not limited to a publication date or study setting but was limited to English-language publications. Any additional relevant studies identified were investigated, and the literature search was updated until January 3, 2018.

Study Selection

We used the following eligibility criteria: 1) patients with biopsy-proved HNSCC who underwent preoperative MR imaging including DWI, 2) histopathology as a reference standard, 3) provision of the diagnostic performance and corresponding ADC cutoff value for differentiating metastatic cervical LNs from benign cervical LNs, and 4) published original articles. Case reports/series (including <10 patients), reviews, conference abstracts, and studies including other types of tumor (including nasopharyngeal carcinoma or lymphoma), or a study population overlapping other studies were excluded. Authors of the studies were contacted for further information when 2 × 2 tables could not be acquired.

Data Extraction and Quality Assessment

The following information was extracted from the selected studies using a standardized form: 1) study characteristics: authors, publication years, institution, study period, study design, and data analysis (per LN versus per patient); 2) demographic characteristics: sample size, mean age, age range, sex, and proportion of metastatic LNs; 3) MR imaging characteristics: magnetic field strength, MR imaging vendor, MR imaging scanner, coil, DWI sequence, b-values (seconds/square millimeter), TR, TE, slice thickness, interslice gap, matrix, FOV, number of signal acquisitions, scan time, number of readers, experience of readers; and 4) outcomes: a 2 × 2 contingency table (number of true-positive, false-positive, false-negative, and true-negative results) demonstrating the presence of metastatic LNs according to the ADC values and optimal ADC cutoff values for differentiating metastatic from benign LNs.

The risk of bias was assessed for each selected study, according to the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) criteria.[16] Two reviewers (C.H.S. and Y.J.C.) independently performed study selection, data extraction, and quality assessment.

Statistical Analysis

A diagnostic meta-analysis of the DWI was conducted with a bivariate random-effects model.[17-19] Individual study sensitivity/specificity and pooled sensitivity/specificity were plotted using a coupled forest plot. The pooled positive likelihood ratio, negative likelihood ratio, and diagnostic odds ratio were calculated. "Positive likelihood ratio" was defined as the likelihood that a DWI result positive for differentiating metastatic LNs from benign LNs would occur in patients with metastatic LNs. "Negative likelihood ratio" was defined as the likelihood that a DWI result negative for differentiating metastatic LNs from benign LNs would occur in patients without metastatic LNs. The "diagnostic odds ratio" was defined as the odds of having a positive DWI result in patients with metastatic LNs compared with the odds of having a positive DWI result in patients without metastatic LNs. A hierarchical summary receiver operating characteristic curve with 95% confidence and prediction regions was obtained, and the area under the hierarchical summary receiver operating characteristic curve was calculated.

Heterogeneity across the studies was explored using the inconsistency index (I^2) and Cochran Q-statistics.[20] I^2 values of >50% indicated the presence of heterogeneity across the studies.[21] Visual assessment of a coupled forest plot (inverse correlation indicating the presence of a threshold effect) and a Spearman correlation (a coefficient of >0.6 indicating the presence of a threshold effect) were performed to evaluate any threshold effect (positive correlation between sensitivity and the false-positive rate).[22] Visual assessment of the difference between the 95% confidence and prediction regions in the hierarchical summary receiver operating characteristic curve (a large difference indicating heterogeneity) was also performed. The presence of publication bias was assessed by a Deeks funnel plot asymmetry test,[23] and a slope coefficient with $P < .1$ was considered significant small-study bias.

Meta-regression was performed using the following covariates: 1) analysis method (per LN versus per patient); 2) the percentage of metastatic LNs (<36.7% [median value of the included studies] versus ≥36.7%); 3) underlying disease (HNSCC versus oral squamous cell carcinoma); 4) study design (prospective versus retrospective); 5) consecutive enrollment; 6) number of readers (2 versus 1); 7) magnetic field strength (3T versus 1.5T); 8) maximum b-value (<1000 versus 1000 s/mm^2); 9) slice thickness (3 versus >3 mm); and 10) ADC measurement (whole node versus single section).

All statistical analyses were conducted by one of the reviewers (C.H.S., with 5 years of experience in conducting systematic reviews and meta-analyses) using commercially available software (STATA 15.0, StataCorp, College Station, Texas; and R statistical and computing software, Version 3.4.1; http://www.r-project.org/). $P < .05$ indicated statistical significance.

RESULTS

Study Search

Figure 1 provides an overview of the search strategy and study-selection procedure. After 15 non-English studies were excluded, our search yielded 214 records, of which 35 articles remained after screening of the titles and abstracts. The full text of these studies was reviewed, and 26 studies were excluded as follows (On-line Appendix): studies evaluating patients with enlarged cervical LNs (not all patients had HNSCC [n = 9]), studies including patients with non-HNSCC malignancy (nasopharyngeal carcinoma or lymphoma [n = 5]), a study population partially overlapping other studies (n = 4), studies that did not allow a 2 × 2 contingency table to be obtained (n = 4), and studies not in the field of interest (n = 4). There were no studies reporting the diagnostic performance of DWI without mentioning the corresponding ADC cutoff value. Ultimately, 9 studies with 337 patients were included in this meta-analysis and were considered for further analyses.[4-12]

Study Characteristics and Quality Assessment

The relevant study characteristics are summarized in On-line Table 1. Eight of 9 studies analyzed the diagnostic performance of DWI per LN,[4,5,7-12] whereas 1 study performed analysis on a per-patient basis.[6] The number of included patients ranged from 16 to 80, and the number of LNs ranged from 34 to 651. There were 7 prospective studies[4-8,10,11] and 1 retrospective one,[9] with the study design not being explicit in a further study.[12] Informed consent was obtained in 8 studies,[4,6-12] as was approval by an ethics committee or institutional review board.[4-8,10-12]

The results of the methodologic quality assessment according to QUADAS-2 are presented in Fig 2. Most studies were considered to have a low risk of bias and minimal concerns regarding applicability. Common weaknesses involved uncertainties in blinding to the reference standard when analyzing the MR imaging results and a poorly documented time interval between MR imaging and the reference standard. In the patient-selection domain, 2 studies had a high risk of bias due to a case-control design[9] or inappropriate exclusion criteria.[8] In the index test domain, 3 studies had an unclear risk of bias because no information was provided on blinding to the reference standard.[6,7,12] In the reference standard and flow/timing domain, 1 study had a high risk of bias and a high concern regarding applicability because both histopathology and follow-up imaging results were used as a reference standard.[9] No studies were excluded from the meta-analysis on the basis of the quality assessment.
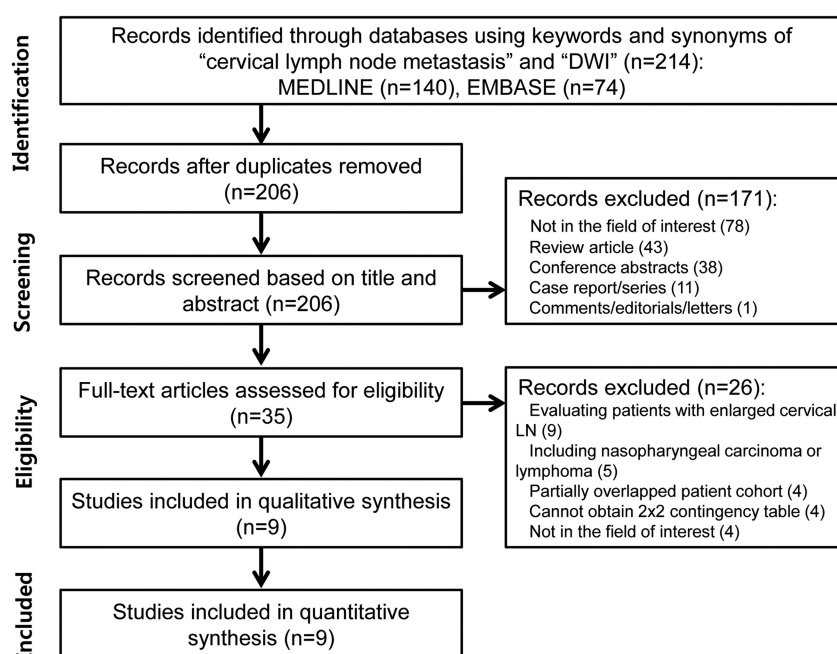

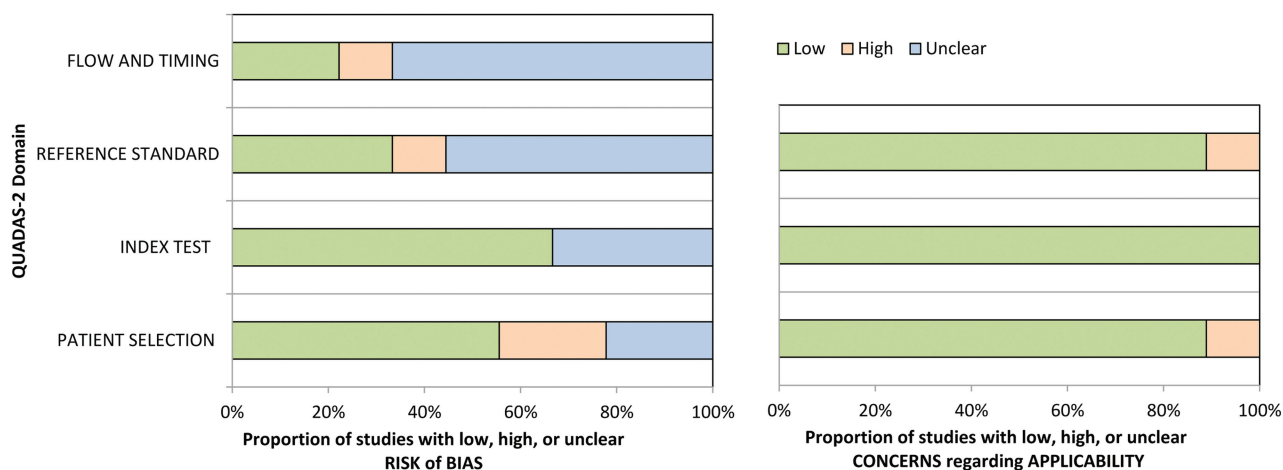**FIG 1.** Flowchart depicting the literature search and study selection.

FIG 2. Grouped bar charts indicating methodologic quality according to the QUADAS-2 criteria and expressed as the percentage of studies meeting each criterion. For each quality domain, the proportions of studies suggesting a low, high, or unclear risk of bias and/or concerns regarding applicability are illustrated in green, red, and blue, respectively.
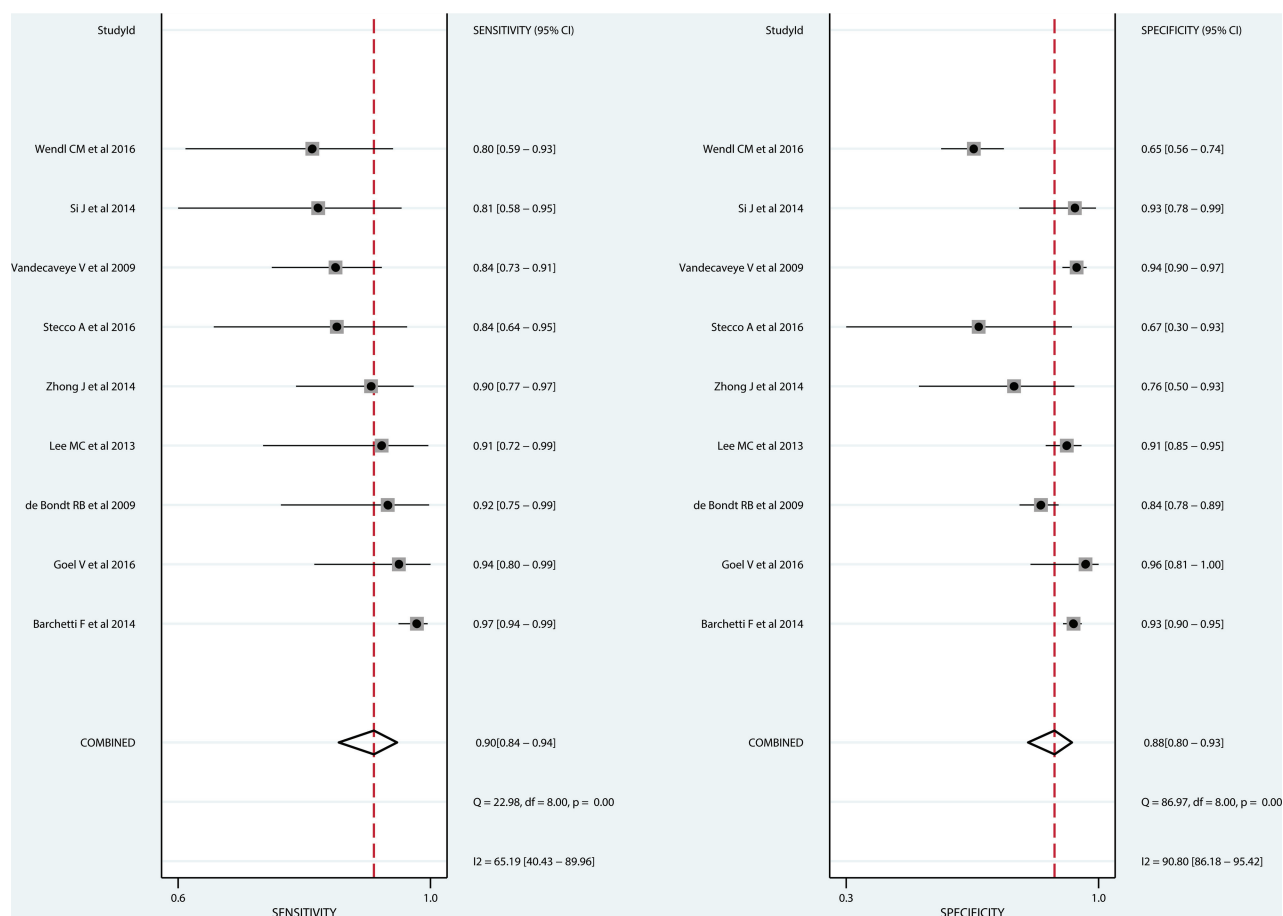
FIG 3. Coupled forest plots showing the diagnostic performance of each study. *Vertical lines* in the coupled forest plots show the pooled sensitivity and specificity.

MR Imaging Characteristics

The detailed MR imaging parameters are summarized in On-line Table 2. Four studies used 3T,[4,7,8,11] and 5 studies used 1.5T[5,6,9,10,12] MR imaging. Six studies used echo-planar imaging as a DWI sequence,[4,5,7,8,10,11] with the other studies not being explicit.[6,9,12] Two b-values were used for DWI in 4 studies[4,5,7,12]; 3 b-values, in 4 studies[6,8,9,11]; and 6 b-values, in 1 study.[10] Four studies used a 3-mm slice thickness,[4,5,7,11] and 4 studies used a 4- or 5-mm slice thickness.[6,8-10]

ROIs were drawn manually around each LN in all studies, with the ROIs being placed on solid portions and avoiding cystic or necrotic portions. An ADC value covering the whole node was obtained in 5 studies,[4-6,8,9] while an ADC value for a single section was obtained in 4 studies.[7,10-12] The mean ADC was calculated in 8 studies,[4-8,10-12] and the minimum ADC (the lowest value) was calculated in 1 study.[9] The smallest LN sizes for ADC calculation were set at a minimal axial diameter of 2 mm,[5,7] 3 mm,[11] 4 mm,[4] or 10 mm.[8]

Data Analysis

In all studies, ADC values derived from metastatic LNs were significantly lower than ADC values derived from benign LNs. The optimal ADC cutoff values varied slightly among individual studies, ranging from 0.851×10^{-3} mm^2/s to 1.038×10^{-3} mm^2/s. The median ADC cutoff value was 0.965×10^{-3} mm^2/s. The

individual sensitivities ranged from 80% to 97%, and the individual specificities ranged from 65% to 96%.

The pooled sensitivity and specificity for the diagnostic performance of DWI in differentiating metastatic LNs from benign cervical LNs were 90% (95% CI, 84%–94%) and 88% (95% CI, 80%–93%), respectively (Fig 3). The pooled positive likelihood ratio, negative likelihood ratio, and diagnostic odds ratio were 7.3 (95% CI, 4.4–12.1), 0.11 (95% CI, 0.07–0.19), and 64 (95% CI, 27–156), respectively. The area under the hierarchical summary receiver operating characteristic curve was 0.95 (95% CI, 0.93–0.97), which suggests high diagnostic performance (Fig 4).

Heterogeneity was present, with I^2 values exceeding 50% for both sensitivity and specificity. Visual assessment of the coupled forest plots revealed no threshold effect, and the Spearman correlation coefficient was −0.471 (95% CI, −0.865–0.281), also indicating no threshold effect. The slope coefficient for the Deeks funnel plot for differentiating metastatic LNs from benign LNs is presented in Fig 5 and suggests slight asymmetry in the data (*P* = .03) and possible publication bias.

The Table shows the results of the meta-regression to explore the influence of 10 covariates on pooled sensitivity and specificity. Slice thickness was revealed to be a significant factor affecting study heterogeneity. Sensitivity was significantly higher in studies
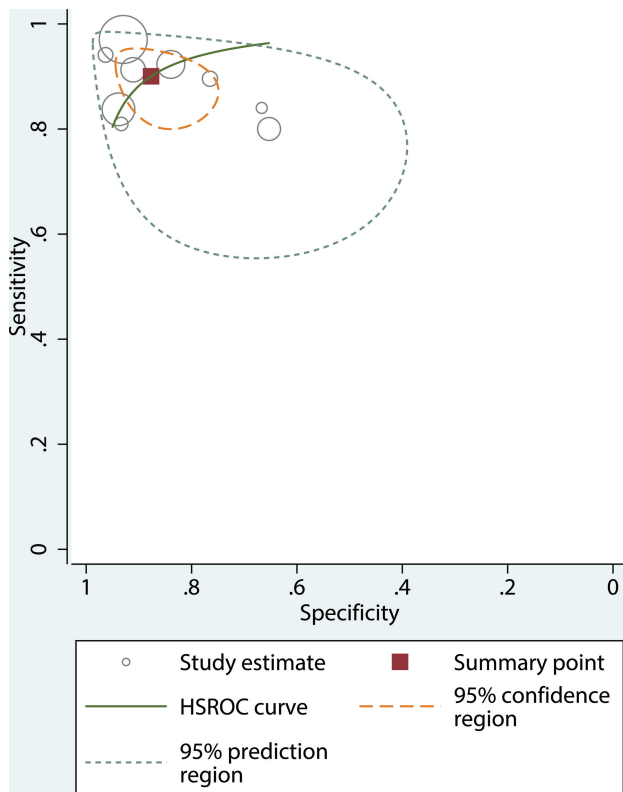
FIG 4. Hierarchical summary receiver operating characteristic curve with 95% confidence and prediction regions of DWI for differentiating metastatic LNs from benign LNs in patients with HNSCC. Each *circle* indicates 1 included study.
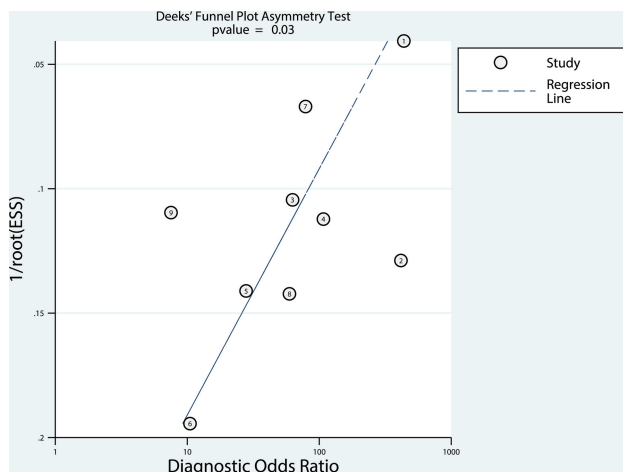

FIG 5. The Deeks funnel plot showing the presence of publication bias. Numbers in *circles* refer to study number. ESS indicates effective sample size.

using a 3-mm slice thickness (93% [95% CI, 88%–98%]) than in studies using a slice thickness of >3 mm (86% [95% CI, 77%–95%], $P < .01$). Otherwise, the analysis method, percentage of metastatic LNs, underlying disease, study design, consecutive enrollment, number of readers, magnetic field strength, maximum b-value, and ROI used for ADC measurement were not significant factors affecting heterogeneity. MR imaging using a 3T scanner, a maximum b-value of 1000 s/mm², and an ADC measurement of the whole node all

showed slightly higher sensitivity; however, the differences did not reach statistical significance.

DISCUSSION

The present systematic review and meta-analysis demonstrated that in patients with HNSCC, the ADC derived from metastatic LNs was significantly lower than the ADC derived from benign LNs. The median ADC cutoff value was 0.965×10^{-3} mm²/s. In addition, our study demonstrated a high pooled sensitivity and specificity for the diagnostic performance of DWI for differentiating metastatic from benign LNs in patients with HNSCC. The meta-regression revealed that studies using a 3-mm slice thickness had higher sensitivity than studies using a slice thickness of >3 mm. Therefore, DWI using a 3-mm slice thickness should be optimally considered for differentiating metastatic from benign cervical LNs.

A dedicated sequence optimization is essential to obtain optimized DWI. In the meta-regression, sensitivity was significantly higher in studies using a 3-mm slice thickness (93% [95% CI, 88%–98%]) than in studies using a slice thickness of >3 mm (86% [95% CI, 77%–95%], $P < .01$). A 3-mm slice thickness may help to detect smaller sized LNs. In addition, 3T MR imaging (92% [95% CI, 86%–98%]), the use of a maximum b-value of 1000 s/mm² (91% [95% CI, 86%–97%]), and ADC measurement of the whole node (93% [95% CI, 89–97]) showed slightly higher sensitivity, though the differences did not reach statistical significance. A previous report also mentioned that a high gradient strength substantially increases the signal-to-noise ratio and that applying a larger number of b-values not only reduces the influence of noise propagation in ADC calculations but also decreases the risk of motion-related artifacts.[10] When one uses DWI to differentiate metastatic from benign LNs in patients with HNSCC, use of a 3-mm slice thickness, a 3T scanner, a maximum b-value of 1000 s/mm², and ADC measurement of the whole node should all be considered to obtain a high diagnostic performance. Considerable effort is required to achieve standardization, and further studies are needed.

Among the included studies, the ADCs derived from metastatic LNs were consistently lower than the ADCs derived from benign LNs. The lower ADC values are probably due to the tumor microstructure in metastatic LNs, which typically show a larger number of cells, cellular polymorphism, and increased mitosis in comparison with benign LNs; these characteristics may reduce the extracellular extravascular space and decrease the ADC value.[24] In our study, the optimal ADC cutoff values ranged from 0.851×10^{-3} mm²/s to 1.038×10^{-3} mm²/s, with 7 of 9 studies reporting optimal cutoff values between 0.94×10^{-3} mm²/s and 1.038×10^{-3} mm²/s, which is a relatively small variation. In addition, the median ADC cutoff value was 0.965×10^{-3} mm²/s.

We recognize that our study has several limitations. First, a small number of studies were included in our meta-analysis; therefore, we cannot evaluate all potential causes of heterogeneity. Although we found that slice thickness was a significant factor affecting study heterogeneity, other technical aspects, including different TRs/TEs and different sets of b-values, may

Metaregression results

Covariates	Subgroups	Pooled Sensitivity (95% CI)	P Value	Pooled Specificity (95% CI)	P Value
Analysis method	Per LN	89% (84%–94%)	.42	87% (80–93%)	.43
	Per patient	95% (85%–100%)		97% (89–100%)	
Percentage of metastatic LN	<36.7%	86% (78%–94%)	.20	86% (77–95%)	.25
	≥36.7%	92% (88%–97%)		89% (81%–97%)	
Underlying disease	HNSCC	91% (86%–96%)	.29	88% (81%–95%)	.37
	Oral SCC	87% (76%–97%)		86% (74%–99%)	
Study design	Prospective	91% (86%–96%)	.74	89% (84%–94%)	.15
	Retrospective	85% (64%–100%)		68% (24%–100%)	
Consecutive enrollment	Yes	92% (87%–96%)	.30	91% (89%–94%)	.92
	No	86% (76%–96%)		68% (56%–80%)	
No. of readers	2	90% (84%–95%)	.42	87% (80%–94%)	.53
	1	89% (76%–100%)		85% (70%–100%)	
Magnetic field strength	3T	92% (86%–98%)	.12	87% (79%–96%)	.11
	1.5T	89% (82%–95%)		88% (80%–96%)	
Maximum b-value	<1000	88% (79%–96%)	.16	86% (76%–97%)	.27
	1000	91% (86%–97%)		89% (82%–96%)	
Slice thickness	3 mm	93% (88%–98%)	<.01	86% (78%–94%)	.49
	>3 mm	86% (77%–95%)		93% (88%–99%)	
ADC measurement	Whole node	93% (89%–97%)	.19	89% (82%–97%)	.28
	Single section	86% (79%–94%)		86% (76%–95%)	

Note:—ADC indicates apparent diffusion coefficient; HNSCC, head and neck squamous cell carcinoma; LN, lymph node; SCC, squamous cell carcinoma.

account for some portion of the heterogeneity. In addition, the low number of included studies may limit the power to achieve statistical significance. Second, publication bias was reported. One possible reason is that 2 studies showing negative results were excluded because of the nonavailability of 2 × 2 contingency tables.[13,14] Therefore, our results should be interpreted cautiously, and the high diagnostic performance of DWI may have been overestimated. To overcome these limitations, we included a relatively homogeneous study population (ie, biopsy-proved HNSCC) and performed an extensive meta-regression using 10 covariates. Moreover, we applied recent robust methodology (hierarchical logistic regression modeling[17-19]) and reported our results according to prestigious guidelines (the Preferred Reporting Items for Systematic Review and Meta-Analysis[15] and the *Handbook for Diagnostic Test Accuracy Reviews* published by the Cochrane Collaboration[25]). Nevertheless, caution should be used when applying our results to daily clinical practice.

CONCLUSIONS

DWI demonstrated a high diagnostic performance for differentiating metastatic from benign cervical LNs in patients with HNSCC, and the median ADC cutoff value was 0.965×10^{-3} mm^2/s. A 3-mm slice thickness for DWI can slightly improve sensitivity. Further large prospective multicenter studies are required to confirm these findings.

REFERENCES

1. NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines®) for Head and Neck Cancers V.2.2017. https://www.nccn.org/professionals/physician_gls/pdf/head-and-neck.pdf. Accessed January 23, 2018. To view the most recent and complete version of the guidelines, National Comprehensive Cancer Network®, NCCN®, NCCN Guidelines®, and all other NCCN content that are trademarks owned by the National Comprehensive Cancer Network, Inc, see www.nccn.org

2. Curtin HD, Ishwaran H, Mancuso AA, et al. **Comparison of CT and MR imaging in staging of neck metastases.** *Radiology* 1998;207:123–30 CrossRef Medline

3. van den Brekel MW, Castelijns JA, Snow GB. **Detection of lymph node metastases in the neck: radiologic criteria.** *Radiology* 1994;192:617–18 CrossRef Medline

4. Barchetti F, Pranno N, Giraldi G, et al. **The role of 3 Tesla diffusion-weighted imaging in the differential diagnosis of benign versus malignant cervical lymph nodes in patients with head and neck squamous cell carcinoma.** *Biomed Res Int* 2014;2014:532095 CrossRef Medline

5. de Bondt RB, Hoeberigs MC, Nelemans PJ, et al. **Diagnostic accuracy and additional value of diffusion-weighted imaging for discrimination of malignant cervical lymph nodes in head and neck squamous cell carcinoma.** *Neuroradiology* 2009;51:183–92 CrossRef Medline

6. Goel V, Parihar PS, Parihar A, et al. **Accuracy of MRI in prediction of tumour thickness and nodal stage in oral tongue and gingivobuccal cancer with clinical correlation and staging.** *J Clin Diagn Res* 2016;10:Tc01–05 CrossRef Medline

7. Lee MC, Tsai HY, Chuang KS, et al. **Prediction of nodal metastasis in head and neck cancer using a 3T MRI ADC map.** *AJNR Am J Neuroradiol* 2013;34:864–69 CrossRef Medline

8. Si J, Huang S, Shi H, et al. **Usefulness of 3T diffusion-weighted MRI for discrimination of reactive and metastatic cervical lymph nodes in patients with oral squamous cell carcinoma: a pilot study.** *Dentomaxillofac Radiol* 2014;43:20130202 CrossRef Medline

9. Stecco A, Buemi F, Cassarà A, et al. **Comparison of retrospective PET and MRI-DWI (PET/MRI-DWI) image fusion with PET/CT and MRI-DWI in detection of cervical and endometrial cancer lymph node metastases.** *Radiol Med* 2016;121:537–45 CrossRef Medline

10. Vandecaveye V, De Keyzer F, Vander Poorten V, et al. **Head and neck squamous cell carcinoma: value of diffusion-weighted MR imaging for nodal staging.** *Radiology* 2009;251:134–46 CrossRef Medline

11. Wendl CM, Müller S, Eiglsperger J, et al. **Diffusion-weighted imaging in oral squamous cell carcinoma using 3 Tesla MRI: is there a chance for preoperative discrimination between benign and malignant lymph nodes in daily clinical routine?** *Acta Radiol* 2016;57:939–46 CrossRef Medline

12. Zhong J, Lu Z, Xu L, et al. **The diagnostic value of cervical lymph node metastasis in head and neck squamous carcinoma by using**

diffusion-weighted magnetic resonance imaging and computed tomography perfusion. *Biomed Res Int* 2014;2014:260859 CrossRef Medline

13. Liang L, Luo X, Lian Z, et al. **Lymph node metastasis in head and neck squamous carcinoma: efficacy of intravoxel incoherent motion magnetic resonance imaging for the differential diagnosis.** *Eur J Radiol* 2017;90:159–65 CrossRef Medline

14. Lim HK, Lee JH, Baek HJ, et al. **Is diffusion-weighted MRI useful for differentiation of small non-necrotic cervical lymph nodes in patients with head and neck malignancies?** *Korean J Radiol* 2014;15:810–16 CrossRef Medline

15. 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.** *Ann Intern Med* 2009;151:W65–94 Medline

16. Whiting PF, Rutjes AW, Westwood ME, et al; QUADAS-2 Group. **QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies.** *Ann Intern Med* 2011;155:529–36 CrossRef Medline

17. 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 CrossRef Medline

18. Kim KW, Lee J, Choi SH, et al. **Systematic review and meta-analysis of studies evaluating diagnostic test accuracy: a practical review for clinical researchers, Part I: general guidance and tips.** *Korean J Radiol* 2015;16:1175–87 CrossRef Medline

19. Lee J, Kim KW, Choi SH, et al. **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 CrossRef Medline

20. Higgins JP, Thompson SG. **Quantifying heterogeneity in a meta-analysis.** *Stat Med* 2002;21:1539–58 CrossRef Medline

21. Higgins JP, Thompson SG, Deeks JJ, et al. **Measuring inconsistency in meta-analyses.** *BMJ* 2003;327:557–60 CrossRef Medline

22. Devillé WL, Buntinx F, Bouter LM, et al. **Conducting systematic reviews of diagnostic studies: didactic guidelines.** *BMC Med Res Methodol* 2002;2:9 CrossRef Medline

23. 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 CrossRef Medline

24. Koh DM, Collins DJ. **Diffusion-weighted MRI in the body: applications and challenges in oncology.** *AJR Am J Roentgenol* 2007;188:1622–35 CrossRef Medline

25. Deeks JJ BP, Gatsonis C, eds. **2013 Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy, Version 1.0.0.** The Cochrane Collaboration. http://srdta.cochrane.org/handbook-dta-reviews. Accessed January 8, 2017