J Huazhong Univ Sci Technol［Med Sci］
DOI 10.1007/s11596-014-1379-9
34(6):950-956,2014

Meta-analysis of Quantitative Diffusion-weighted MR Imaging in Differentiating Benign and Malignant Pancreatic Masses*

Xiang-ke NIU (牛翔科)[1], Anup Bhetuwal[2], Sushant Das[2], Ying-quan XIAO (肖应权)[2], Feng SUN (孙 凤)[2], Li-chuan ZENG (曾利川)[2], Han-feng YANG (杨汉丰)[2#]

[1]*Department of Radiology, Affiliated Hospital of Chengdu University, Chengdu 610000, China*
[2]*Sichuan Key Laboratory of Medical Imaging and Department of Radiology, Affiliated Hospital of North Sichuan Medical College, Nanchong 637000, China*

Summary: There have been numerous studies done to explore the diagnostic performance of quantitative diffusion-weighted (DW) MR imaging to differentiate between benign and malignant pancreatic masses. However, the results have been inconsistent. We performed a meta-analysis to investigate whether DW-MR imaging can differentiate between these two diseases. Databases including MEDLINE, EMBASE and Cochrane Library were utilized to find relevant articles published between January 2001 and January 2014. A Stata version 12.0 and a Meta-Disc version 1.4 were used to describe primary results. Twelve studies with 594 patients, which fulfilled the inclusion criteria, were enrolled for the analysis. The pooled sensitivity and specificity of DW imaging was 0.91 (95% CI: 0.84, 0.95) and 0.86 (95% CI: 0.76, 0.93) respectively. The area under the curve of the summary receiver operating characteristic was 0.95 (95% CI: 0.93, 0.96). The results indicated that DW imaging might be a valuable tool for differentiating benign and malignant pancreatic masses.
Key words: apparent diffusion coefficient; diffusion-weighted imaging; magnetic resonance imaging; meta-analysis; pancreatic tumor; pancreatic adenocarcinoma

Pancreatic cancer, which represents a major diagnostic and therapeutic concern, is characterized by high mortality and short survival time even after early diagnosis. Current imaging methods can achieve a radiological diagnosis of high sensitivity and specificity only in advanced tumors, often with only palliative therapeutic options left[1]. Despite the great technical advances in imaging, such as ultrasonography, multidetector computed tomography (MDCT) and magnetic resonance imaging (MRI), differentiation between cancerous and noncancerous pancreatic masses at an early stage is not satisfactory and new methods are justifiably required[2].

The principle of Brownian motion (random thermal diffusion) of small molecules in a tissue is the basis of diffusion-weighted imaging (DWI). The important role of DWI in oncological applications is widely accepted and its value for abdominal lesion detection and characterization is increasingly investigated[3]. The implementation of ultrafast imaging techniques, such as parallel imaging, has made DWI of the upper abdomen a feasible option. This has been found to be useful in the differentiation of malignant from benign liver lesions[4, 5].

However, while the diagnostic performances of pancreatic quantitative DWI were examined in numerous studies, the differences in the patient characteristics, MR imaging techniques and diagnostic criteria for malignancy cause inconsistent results[6]. The aim of this study was to review published studies that utilized DWI to detect malignant and benign pancreatic lesions, and to evaluate the overall diagnostic value of DWI in the differentiation of pancreatic lesions using a meta-analysis. This would allow us to establish a non-invasive imaging protocol in the clinical routine to differentiate between cancerous and noncancerous pancreatic masses.

1 MATERIALS AND METHODS

This meta-analytic review was done with reference to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement[7, 8].

1.1 Literature Search

We performed a comprehensive computer literature search to identify articles on the diagnostic performance of DWI in the differentiation of pancreatic lesions published from January 2001 to January 2014 in the MEDLINE, EMBASE and Cochrane Library databases. A search was performed with the following terms or MeSH subject headings used: "diffusion-weighted magnetic resonance images" OR "diffusion magnetic resonance" OR "DW-MRI" OR "DW magnetic resonance images" AND "pancreas" AND "pancreatic tumor" OR "pancreatic adenocarcinoma". A manual search of list of references included in the studies and review articles was done.

Xiang-ke NIU, E-mail: 862639224@qq.com
#Corresponding author, E-mail: niu19850519@163.com
*This project was supported by the Science Foundation for Distinguished Young Scholars of Sichuan Province, China (No. 2012JQ0060).

1.2 Selection of Articles

Two investigators, blinded to the journal, author, institution and date of publication, independently checked the retrieved articles. The two investigators resolved any type of disagreement by consensus. By using standardized data extraction form, we read each abstract to procure the potentially eligible articles. Then we retrieved the full text to determine whether the articles were eligible for our study or not. The inclusion criteria for the articles were (a) article published in English; (b) use of DWI to differentiate between cancerous and noncancerous pancreatic masses; (c) sufficient information to calculate true-positive (TP), false-negative (FN), false-positive (FP), and true-negative (TN) values for per patient-level statistics; (d) total number of lesions > 20; (e) assessing the quality of study design. If the number of "Yes" answers to 14 questions was > 9 in the Quality Assessment of Diagnostic Accuracy Studies (QUADAS), a quality assessment tool specifically developed for systematic reviews of diagnostic accuracy studies[9], the article was included. The article was excluded if the number of ''No'' or ''unclear'' answers was > 4; (f) if two articles contained similar data or subsets of data, the one with the most details or most recent publication date was chosen; and (g) studies which applied histo-pathological analysis (performed at surgery or biopsy) and/or clinical and imaging follow-up as a reference standard. We contacted the authors of abstracts and studies with insufficient data to request for additional information regarding their studies.

1.3 Data Extraction and Quality Assessment

The two investigators who selected the eligible articles also performed the relevant data extraction independently. To resolve disagreement between the reviewers, a third reviewer assessed all of the involved items. Two observers assessed the methodological quality of included studies independently by using the QUADAS. We utilized QUADAS criteria again to further examine relevant studies. To perform accuracy analyses, we extracted data on characteristics of studies, patients, measurements and results. For each report, we extracted the following items: author, year of publication, study population, study design (prospective, retrospective or unknown), patient enrollment (consecutive or not), different b values, and magnetic field strength. We calculated the values of TP, FP, FN, and TN for the detection of lesions from each study and constructed 2 × 2 contingency tables.

1.4 Statistical Analysis

We extracted or calculated the sensitivity and specificity of the techniques assessed in a given study using 2 × 2 contingency tables. We also added a value of 0.5 to all cells of studies that had a count of zero to avoid potential problems in odds calculations for studies with sensitivities or specificities of 100%.

1.4.1 Summary Performance Estimates We calculated summary sensitivity and specificity after the antilogit transformation of estimated model parameters. We then derived corresponding positive likelihood, negative likelihood and diagnostic odds ratios as functions of these summary estimates. We also used the derived estimates of sensitivity, specificity, and respective variances to construct a summary receiver operating characteristic (ROC) curve. The area under the ROC curve was then used as an alternative global measurement of test performance. Next was the calculation of the diagnostic odds ratio (DOR). DOR is the odds of having a positive test result among patients with the given disease compared to the odds of having a positive test result among patients without the disease. The method of calculation is as follows: DOR=LR+/LR–, where LR+ is the positive likelihood ratio and LR– is the negative likelihood ratio. This measure is a single indicator of test accuracy that comprises a combination of sensitivity and specificity information. The posttest probability of cancerous pancreatic masses (P_{post}) is calculated from likelihood ratios by using the Bayes theorem as follows: $P_{post} = (LR \times P_{pre})/[(1-P_{pre}) \times (1-LR)]$, where P_{pre}, the pretest probability, is the suspicion for cancerous pancreatic masses.

1.4.2 Homogeneity Test We used the inconsistency index (I-squared, I^2) to estimate the heterogeneity of individual studies contributing to the pooled estimate. I^2 describes the percentage of total variation across studies due to heterogeneity rather than chance. Furthermore, it is used as a measure to quantify the amount of heterogeneity. I^2 >50% suggests heterogeneity[10]. We finalized the pooling of data within the bivariate mixed-effects binary regression-modeling framework. If there were notable heterogeneities, a random-effects model summarized the test performance, otherwise a fixed-effects model was used[11].

1.4.3 Publication Bias Analysis There are more chances of studies with optimistic results for publication than those with unfavorable results. Since publication biases would tend to exaggerate clinical effects, resulting in potentially erroneous clinical decision, it is important to assess the likely extent of the bias and its potential impact on the final conclusion[12]. To assess the publication bias, we utilized the Deeks' Funnel Plot Asymmetry Test. If there occurred a nonzero slope coefficient (P<0.10), we considered a publication bias to be present.

1.4.4 Threshold Effect Threshold effect was one important extra source of variation in meta-analysis. If the threshold effect exists, an inverse correlation appears. Combining the study results concerning fitting of an ROC curve was better than pooling sensitivities and specificities together in these cases. We assessed representation of accuracy estimates from each study in a ROC space and computation results of spearman correlation coefficient between the log (SEN) and log (1–SPE) for any threshold effect. A typical pattern of "shoulder arm" plot in a ROC space along with a strong positive correlation would suggest threshold effect[13, 14].

1.4.5 Subgroup and Sensitivity Analysis Subgroup analyses were also performed according to patient enrollment type (consecutive *vs.* nonconsecutive or unreported), study design (prospective *vs.* retrospective) and b value (b ≥ or <500 s/mm^2). We reappraised the pooled estimates due to exclusion of one study. We then compared the reappraised results with the original results to assess stability and reliability of our meta-analysis.

All the statistical computations were performed

using Stata/SE software (version 12.0, StataCorp) and Meta-Disc (version 1.4, Javier Zamora). *P* values less than 0.05 were considered statistically significant.

2 RESULTS

2.1 Literature Search and Selection of Studies

A total of 118 abstracts were retrieved after the computer search and manual crosschecking of reference lists. Twenty-eight articles were deemed eligible after reading the titles and abstracts. After reading the full texts, we excluded 16 of the 28 relevant articles. The reasons were as follows: (a) not assess the diagnostic value of DWI for pancreatic lesions (*n*=4); (b) insufficient data to calculate TP, FP, TN, and FN (*n*=7); (c) using various diagnostic methods to differentiate pancreatic mass (*n*=2); (d) number of patients < 20 (*n*=2); and (e) presentation of similar data or subsets of data in other articles (*n*=1). Henceforth, twelve studies with 594 patients, with all of the inclusion criteria fulfilled, were enrolled for the analysis. The characteristics of these 12 studies are presented in table 1.

Table 1 The characteristics of included studies

Study authors, year of publication, and reference No.	Number of patients	Study design	Patients' enrollment	b value (s/mm^2)	Field strength (T)	QUADAS score
Kartalis *et al*, 2009[15]	36	Retrospective	Consecutive	0, 500	1.5	11
Conciai *et al*, 2014[16]	33	Retrospective	ND	0, 50	1.5	10
Ichikawa *et al*, 2007[17]	49	ND	ND	0, 1000	1.5	11
Hur *et al*, 2012[18]	36	Retrospective	ND	0, 500	1.5 or 3.0	11
Muhi *et al*, 2012[19]	64	Retrospective	ND	500, 1000	1.5	11
Huang *et al*, 2011[20]	50	ND	Consecutive	0, 1000	3.0	12
Lee *et al*, 2008[21]	60	Prospective	Consecutive	0, 1000	1.5	10
Klauss *et al*, 2011[22]	29	Prospective	ND	0, 200	1.5	13
Kamisawa *et al*, 2010[23]	53	Prospective	Consecutive	0, 50	1.5	11
Sandrasegaran *et al*, 2011[24]	70	Retrospective	ND	50, 400	1.5	11
Fatima *et al*, 2011[25]	69	Retrospective	ND	500, 1000	1.5	10
Schraibman *et al*, 2011[26]	45	Prospective	ND	500, 700	1.5	11

ND, not documented

2.2 Study Description and Assessment of Study Quality

We conducted all analyses based on per-patient data analysis. In these 12 studies, six studies[15, 16, 18, 19, 24, 25] enrolled patients retrospectively whereas four studies[21–23, 26] enrolled patients prospectively. Four studies[15, 20, 21, 23] enrolled patients in a consecutive manner while the others[16–19, 22, 24–26] were not enrolled in a consecutive manner, or the manner of enrollment was unknown. Table 1 shows the principal characteristics of the 12 studies included in the meta-analysis. In general, all studies included in the meta-analysis fulfilled ten or more of the 14 criteria in the QADAS tool for methodological quality.

2.3 Homogeneity Test and Publication Bias

We calculated the I^2 statistic at 15% (I^2=15%), indicating that there was a low heterogeneity among the twelve papers. The next step was the representation of sensitivity against 1–specificity from each study in a ROC space to explore the threshold effect. The pattern of points in this plot did not resemble a "shoulder-arm" shape. We performed a spearman rank correlation test as a further test for threshold effect. The spearman correlation coefficient was equal to –0.600 (*P*=0.208), which indicated that there was no threshold effect in this meta-analysis. The results of Deeks' Funnel Plot Asymmetry Test (*P*=0.696) showed no evidence of notable publication bias (fig. 1).

2.4 Diagnostic Accuracy of DWI

We used a fixed-effects model to calculate diagnostic accuracy. The pooled sensitivity and specificity of DWI was 0.91 (95% CI: 0.84, 0.95) and 0.86 (95% CI: 0.76, 0.93) respectively. The overall AUC was 0.95 (95% CI: 0.93, 0.96), indicating good diagnostic accuracy[27]. Forest plots of the sensitivity and specificity of DWI to discriminate between cancerous and noncancerous pancreatic masses are shown in fig. 2. Summary of ROC curves are shown in fig. 3.

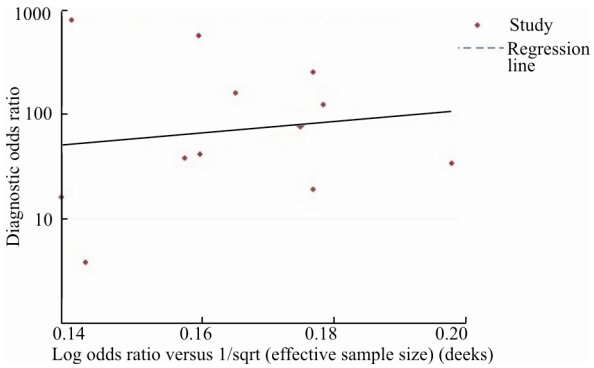

Fig. 1 Results of Deeks' funnel plot of asymmetry test for publication bias

The non-significant slope indicates that no significant bias was found (*P*=0.696).

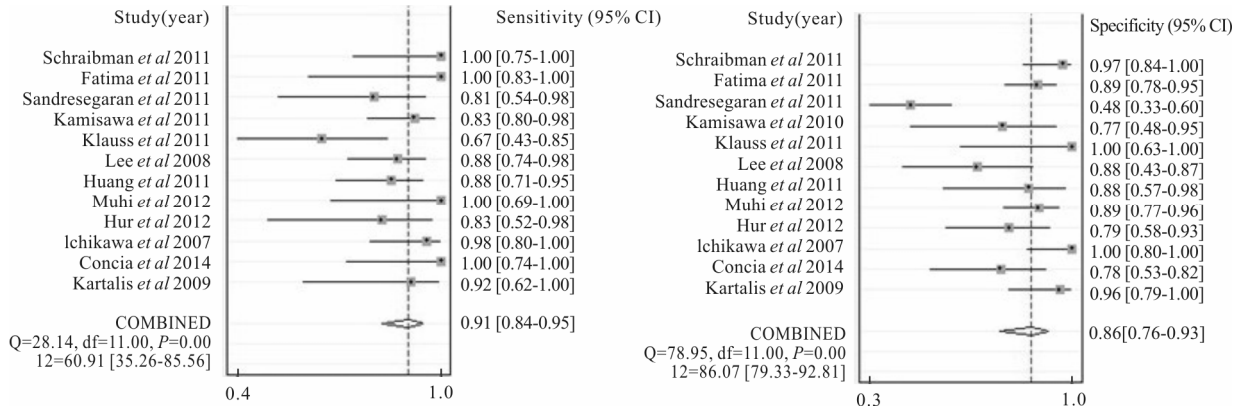

Fig. 2 Forest plot of pooled sensitivity and specificity of DWI in the differentiation of pancreatic lesions
Summary sensitivity and specificity of DWI were 0.91 (95% CI: 0.84, 0.95) and 0.86 (95% CI: 0.76, 0.93), respectively.

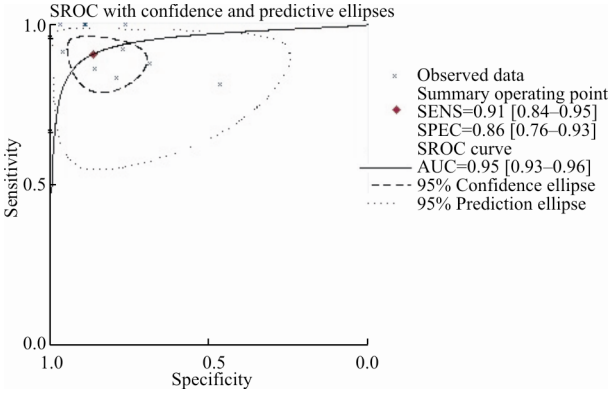

Fig. 3 Summary receiver operating characteristic (sROC) curves of DWI in the differentiation of pancreatic lesions in the 12 included studies.
Numbers in parentheses denote 95% confidence intervals. AUC: area under ROC curve; SENS: sensitivity; SPEC: specificity

2.5 Evaluation of Clinical Utility

The positive and negative likelihood ratios of DWI for differentiating between cancerous and noncancerous pancreatic masses were 6.64 (95% CI: 3.57, 12.32) and 0.11 (95% CI: 0.06, 0.20) respectively. We utilized the likelihood ratios to simulate three clinical scenarios by implementing different pretest probabilities, 25% indicating low clinical suspicion, 75% indicating high clinical suspicion and 50% indicating a worst-case scenario of cancerous pancreatic masses. Using these likelihood ratios, we calculated and plotted the posttest probabilities on Fagan nomograms (fig. 4). With a cancerous lesion's pretest probability of 25% (low clinical suspicion), the posttest probability, given that the DWI result was also negative, dropped to 3%, and whereby we could consider this result sufficient to rule out cancerous lesions (fig. 4A). With a cancerous lesion's pretest probability of 75% (high clinical suspicion), the posttest probability, given that the DWI result is positive, increases to 95%, whereby we could consider this result sufficient to rule in cancerous lesions (fig. 4C). With a pretest probability of cancerous lesions at 50% (worst-case scenario), the posttest probability is 87% if the DWI result was positive, and 10% if the DWI result was negative. Therefore, we can consider DWI to be a useful test in this situation.

We also conducted the subgroup analysis according to the study design (prospective or retrospective), patient enrollment (consecutive or nonconsecutive), and b values (b ≥ or <500 s/mm^2). The results are presented in table 2. With the b value set at ≥ 500 s/mm^2, the pooled data showed higher results for sensitivity (0.93, 95% CI: 0.85, 0.96) and specificity (0.90, 95% CI: 0.81, 0.95) among the studies. The calculated I^2 statistic was 0% (I^2 =0%), which confirmed that there was no evidence of between-study heterogeneity. Omission of any study did not alter the statistical significance of the results (data not shown). Therefore, results of the sensitivity analysis suggested that the data in this meta-analysis were relatively robust.

Table 2 Subgroup analysis

Study characteristics	Reference numbers	Heterogeneity (I^2)	Sensitivity (95% CI)	Specificity (95% CI)	AUC
Study design					
Retrospective	[14, 15, 17, 18, 23, 24]	0%	0.93 (0.81, 0.98)	0.82 (0.67, 0.91)	0.95
Prospective	[20–22, 25]	26.87%	0.87 (0.78, 0.95)	0.91 (0.60, 0.98)	0.93
Patients' enrollment					
Consecutive	[14, 19, 20, 22]	0%	0.89 (0.82, 0.94)	0.84 (0.69, 0.92)	0.90
Nonconsecutive or unclear	[15–18, 21, 23–25]	41.73%	0.94 (0.79, 0.99)	0.88 (0.73, 0.93)	0.97
b value (s/mm^2)					
≥500	[14, 16–20, 24, 25]	0%	0.93 (0.85, 0.96)	0.90 (0.81, 0.95)	0.96
<500	[15, 21, 22, 23]	41.55%	0.87 (0.71, 0.98)	0.76 (0.45, 0.93)	0.90

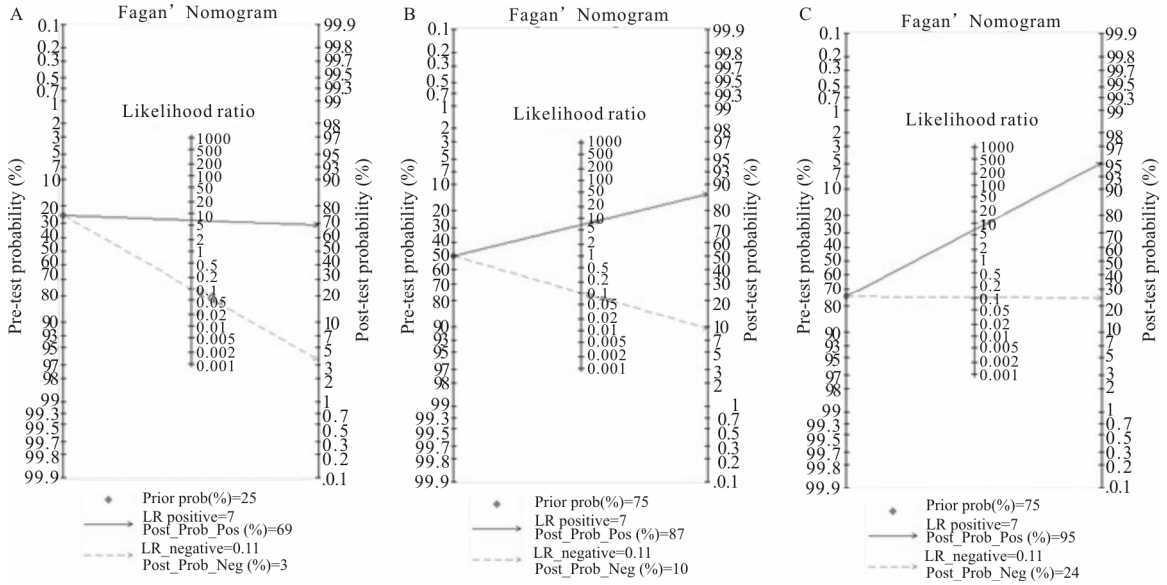

Fig. 4 Pretest probabilities (Prob) and likelihood ratios (LR)

A: With a pretest probability of cancerous lesion at 25% (low clinical suspicion), the post-test probability, with a negative DWI result (Post-Neg Probability), is 3%, which can be considered sufficient to rule out malignancy; B: With a pretest probability of cancerous lesion at 50% (worst-case scenario), the posttest probability, given positive and negative DWI results, is 87% and 10%, respectively; C: With a pretest probability of cancerous lesion at 75% (high clinical suspicion), the posttest probability, given a positive DWI result (Post-Pos Probability), is 95%. Thus, we can consider positive DWI result sufficient to rule in cancerous lesions.

3 DISCUSSION

Multidetector CT (MDCT), positron emission tomography with computed tomography (PET-CT) and endoscopic ultrasound (EUS) have a reported sensitivity of 94%, 89% and 100% and specificity of 87%, 88% and 50%, respectively[28–30]. CT is associated with unavoidable radiation exposure and unpredictable risk of an allergic response to the iodine contrast needed for dynamic analysis. Given an equivalent diagnostic accuracy between CT and MRI, MRI is theoretically, even though it is expensive, an ideal modality for screening. It is well understood that diffusion is caused by random translational molecular motion, also known as brownian water motion. The speed with which water molecules diffuse differs in extracellular and intracellular components of tissues[31]. DWI is the only imaging method that can be used to evaluate the diffusion process *in vivo*.

In this meta-analysis, we calculated an overall sensitivity of 0.91（95% CI: 0.84, 0.95）and specificity of 0.86 (95% CI: 0.76, 0.93) from 12 studies that fulfilled all of the inclusion and exclusion criteria. We found the area under the summary ROC to be 0.95. The results of this systematic review and meta-analysis indicate that DWI may be a useful diagnostic criterion for differentiation between cancerous and noncancerous pancreatic lesions. A previously published meta-analysis for DWI in patients with pancreatic lesions, including 11 studies, reported the overall sensitivity of 0.86 (95% CI: 0.78, 0.91) and specificity of 0.91 (95% CI: 0.81, 0.96). The area under the curve of the summary ROC was 0.94 (95% CI: 0.91, 0.96)[32]. In their study, the heterogeneity in sensitivity tests and specificity tests was highly significant (I^2>60%), confirming that there was strong evidence of between-study heterogeneity[32]. Our meta-analysis calculated I^2 statistic at 15% and as I^2 fell below 25%, there was low heterogeneity among these twelve papers. After careful analysis of study done by Wu *et al*[32], we found that our meta-analyses included literatures different from theirs. It was evident from the study by Wu *et al*[32] that some of the referenced literatures had not accurately differentiated cancerous from non-cancerous pancreatic masses. For example, Inan *et al*[33] showed that DWI had a sensitivity of 70% and specificity of 90% in differentiating abscesses, hydatid cysts and neoplastic cysts from simple cysts and pseudocysts. However, they did not accurately distinguish cancerous lesions from noncancerous pancreatic lesions. Thus, we believe that inclusion of such studies in the study by Wu *et al*[32] is the source of between-study heterogeneity.

The signal intensity of bio-tissues on DWI depends upon the velocity of water molecule diffusion and capillary perfusion. Similarly, this signal is also affected by the T2 shine-through effect since the DWI sequence itself has certain T2 weight[31]. Generally, many studies recommend that a larger b value in DWI be set to reduce interference from the T2 shine-through effect[34]. Ichikawa *et al*[17] and Lee *et al*[21] proposed that DWI with high b value was preferable for differentiating pancreatic mass lesions. In our study, we re-performed a meta-analysis in the subgroup of studies with b≥500 s/mm^2. The results showed that there was no significant heterogeneity. The pooled sensitivity and specificity with 95% CIs were 0.93 (0.85, 0 .96) and 0.90 (0.91, 0.95) respectively. The area under the

curve of sROC was 0.96.

We used Stata and Meta-Disc to assess threshold effect from each study in a ROC space. The spearman correlation coefficient was computed between the log (SEN) and log (1−SPE). Lack of "shoulder-arm" shape of the points in the ROC space and spearman correlation coefficient (0.208) indicated that there should be factors other than differences in cutoff points in DWI to differentiate accuracy estimates between cancerous and noncancerous pancreatic masses across individual studies.

Additionally, publication bias is a usual source of heterogeneity in meta-analysis. In this meta-analysis, we supplemented the search of several electronic databases by checking references of relevant studies in order to reduce publication bias. The funnel plot indicated that there was no significant publication bias in our meta-analysis.

We based our study on thorough literature searches and careful data extraction. It includes assessment of the methodological quality of diagnostic test accuracy studies. Nevertheless, there are some limitations in this meta-analysis. First, our study comprised studies of suboptimal quality. Meta-analysis combined or integrated the results of several independent studies. The quality and reliability of a meta-analysis depends upon the quality of included studies. We utilized the QUADAS tool to assess methodological quality of individual studies. Most included studies in this meta-analysis had a suboptimal design concerning the reporting of time interval between MRI and pathology, absence of differential verification bias, the interpretation of reference standard result without the knowledge of the index test result and the interpretation of the index test result without the knowledge of the reference standard. Second, publication bias is a potential limitation of any meta-analysis. Potential publication bias may perhaps still exist in our study, since it may possibly be easier to publish the studies with optimistic results than those with unfavorable results. Moreover, we only included studies published in English, which might invoke the so-called "Tower of Babel" bias. Third, a selection bias may exist in this study. To avoid the selection bias, we searched for relevant articles in not only the MEDLINE and EMBASE but also Cochrane Library databases. To minimize bias in the selection of studies and in data extraction, reviewers blinded to the journal, author, institution, and date of publication, independently selected articles based on inclusion criteria.

In conclusion, DWI, with pooled sensitivity of 0.91, specificity of 0.86 and area under curve of sROC at 0.95, can be a useful imaging modality to differentiate malignant from benign pancreatic lesions. However, even when all of the methodological issues are considered, we must interpret these results with caution. Large-scale high-quality trials are necessary to assess and confirm the clinical value of DWI. An update of this analysis must be conducted when additional data becomes available.

Conflict of Interest Statement

The authors declare that there is no conflict of interest with any financial organization or corporation or individual that can inappropriately influence this work.

REFERENCES

1 Nanashima A, Tobinaga S, Abo T, *et al*. Evaluation of surgical resection for pancreatic carcinoma at a Japanese single cancer institute. Hepatogastroenterology, 2012, 59(115):911-915

2 Li D, Xie K, Wolff R, *et al.* Pancreatic cancer. Lancet, 2004,363(9414):1049-1057

3 Padhani AR, Liu G, Koh DM, *et al*. Diffusion-weighted magnetic resonance imaging as a cancer biomarker: consensus and recommendations. Neoplasia, 2009,11(2): 102-125

4 Bruegel M, Holzapfel K, Gaa J, *et al*. Characterization of focal liver lesions by ADC measurements using a respiratory triggered diffusion-weighted single-shot echo-planar MR imaging technique. Eur Radiol, 2008,18(3):477-485

5 Gourtsoyianni S, Papanikolaou N, Yarmenitis S, *et al*. Respiratory gated diffusion-weighted imaging of the liver: value of apparent diffusion coefficient measurements in the differentiation between most commonly encountered benign and malignant focalliver lesions. Eur Radiol, 2008,18(3):486-492

6 Kang KM, Lee JM, Yoon JH, *et al*. Intravoxel incoherent motion diffusion-weighted MR imaging for characterization of focal pancreatic lesions. Radiology, 2014,270(2): 444-453

7 Moher D, Liberati A, Tetzlaff J, *et al*. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Int J Surg, 2010,8(5):336-341

8 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(10):e1-e34

9 Whiting P, Rutjes AW, Reitsma JB, *et al*. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol, 2003,3:25

10 Huedo-Medina TB, Sánchez-Meca J, Marín-Martínez F, *et al*. Assessing heterogeneity in meta-analysis: Q statistic or I^2 index? Psychol Methods, 2006,11(2): 193-206

11 Leeflang MM, Deeks JJ, Gatsonis C, *et al.* Cochrane Diagnostic Test Accuracy Working Group. Systematic reviews of diagnostic test accuracy. Ann Intern Med, 2008,149(12):889-897

12 Song F, Khan KS, Dinnes J, *et al*. Asymmetric funnel plots and publication bias in meta-analyses of diagnostic accuracy. Int J Epidemiol, 2002,31(1):88-95

13 Zamora J, Abraira V, Muriel A, *et al*. Meta-DiSc: a software for meta-analysis of test accuracy data. BMC Med Res Methodol, 2006,6:31

14 Moses LE, Shapiro D, Littenberg B. Combining independent studies of a diagnostic test into a summary ROC curve: data-analytic approaches and some additional considerations. Stat Med, 1993,12(14): 1293-1316

15 Kartalis N, Lindholm TL, Aspelin P, *et al*. Diffusion-weighted magnetic resonance imaging of pancreas tumours. Eur Radiol, 2009,19(8):1981-1990

16 Concia M, Sprinkart AM, Penner AH, *et al*. Diffusion-weighted magnetic resonance imaging of the pancreas: diagnostic benefit from an intravoxel incoherent motion model-based 3 b-value analysis. Invest Radiol, 2014,49(2):93-100

17 Ichikawa T, Erturk SM, Motosugi U, et al. High-b value diffusion-weighted MRI for detecting pancreatic adenocarcinoma: preliminary results. AJR Am J Roentgenol, 2007,188(2):409-414
18 Hur BY, Lee JM, Lee JE, *et al*. Magnetic resonance imaging findings of the mass-forming type of autoimmune pancreatitis: comparison with pancreatic adenocarcinoma. J Magn Reson Imaging, 2012,36(1):188-197
19 Muhi A, Ichikawa T, Motosugi U, *et al*. Mass-forming autoimmune pancreatitis and pancreatic carcinoma: differential diagnosis on the basis of computed tomography and magnetic resonance cholangio- pancreatography, and diffusion-weighted imaging findings. J Magn Reson Imaging, 2012,35(4): 827-836
20 Huang WC, Sheng J, Chen SY, *et al*. Differentiation between pancreatic carcinoma and mass-forming chronic pancreatitis: usefulness of high b value diffusion-weighted imaging. J Dig Dis, 2011,12(5):401-408
21 Lee SS, Byun JH, Park BJ, *et al*. Quantitative analysis of diffusion-weighted magnetic resonance imaging of the pancreas: usefulness in characterizing solid pancreatic masses. J Magn Reson Imaging, 2008,28(4):928-936
22 Klauss M, Lemke A, Grünberg K, *et al*. Intravoxel incoherent motion MRI for the differentiation between mass forming chronic pancreatitis and pancreatic carcinoma. Invest Radiol, 2011,46(1):57-63
23 Kamisawa T, Takuma K, Anjiki H, *et al*. Differentiation of autoimmune pancreatitis from pancreatic cancer by diffusion-weighted MRI. Am J Gastroenterol, 2010, 105(8):1870-1875
24 Sandrasegaran K, Akisik FM, Patel AA, *et al*. Diffusion-weighted imaging in characterization of cystic pancreatic lesions. Clin Radiol, 2011,66(9):808-814
25 Fatima Z, Ichikawa T, Motosugi U, *et al*. Magnetic resonance diffusion-weighted imaging in the characterization of pancreatic mucinous cysticlesions. Clin Radiol, 2011,66(4):108-111
26 Schraibman V, Goldman SM, Ardengh JC, *et al*. New trends in diffusion-weighted magnetic resonance imaging as a tool in differentiation of serous- cystadenoma and mucinous cystic tumor: a prospective study. Pancreatology, 2011,11(1):43-51
27 Jones CM, Athanasiou T. Summary receiver operating characteristic curve analysis techniques in the evaluation of diagnostic tests. Ann Thorac Surg, 2005,79(1):16-20
28 Fletcher JG, Wiersema MJ, Farrell MA, *et al*. Pancreatic malignancy: value of arterial, pancreatic, and hepatic phase imaging with multi-detectorrow CT. Radiology, 2003,229(1):81-90
29 Farma JM, Santillan AA, Melis M, *et al*. PET/CT fusion scan enhances CT staging in patients with pancreatic neoplasms. Ann Surg Oncol, 2008,15(9):2465-2471
30 Agarwal B, Abu-Hamda E, Molke KL, *et al*. Endoscopic ultrasound-guided fine needle aspiration and multidetector spiral CT in the diagnosis ofpancreatic cancer. Am J Gastroenterol, 2004,99(5):844-850
31 Roth Y, Tichler T, Kostenich G, *et al*. High-b-value diffusion-weighted MR imaging for pretreatment prediction and early monitoring of tumorresponse to therapy in mice. Radiology, 2004,232(3):685-692
32 Wu LM, Xu JR, Hua J, *et al*. Value of diffusion-weighted imaging for the discrimination of pancreatic lesions: a meta-analysis. Eur J Gastroenterol Hepatol, 2012,24(2): 134-142
33 Inan N, Arslan A, Akansel G, *et al*. Diffusion-weighted imaging in the differential diagnosis of cystic lesions of the pancreas. AJR Am J Roentgenol, 2008,191(4): 1115-1121
34 Tajima Y, Kuroki T, Tsutsumi R, *et al*. Pancreatic carcinoma coexisting with chronic pancreatitis versus tumor-forming pancreatitis:diagnostic utility of the time-signal intensity curve from dynamic contrast- enhanced MR imaging. World J Gastroenterol, 2007,13(6): 858-865

(Received Apr. 29, 2014; revised Jul. 14, 2014)