Original Article

Meta-analysis of diffusion-weighted magnetic resonance imaging in identification of colorectal cancer

Hongyuan Jia1*, Xuelei Ma1,2*, Yang Zhao1*, Jingyi Zhao1*, Rongjun Liu1*, Zihang Chen1, Jinna Chen1, Jingwen Huang2*, Yanyan Li1*, Jing Zhang1, Feng Wang1,2

1Cancer Center, State Key Laboratory of Biotherapy, West China Hospital, Sichuan University, Chengdu 610041, PR China; 2Department of Oncology, West China Hospital, Sichuan University, Chengdu 610041, PR China. *Equal contributors.

Received July 15, 2015; Accepted October 12, 2015; Epub October 15, 2015; Published October 30, 2015

Abstract: Purpose: This meta-analysis aimed to evaluate the performance of diffusion-weighted magnetic resonance imaging (DWI) in identification of colorectal cancer. Methods: A systematic literature search was performed for studies that evaluated the diagnostic accuracy of DWI in identification of colorectal cancer. Methodological quality was assessed by Quality Assessment for Studies of Diagnostic Accuracy 2 (QUADAS 2) tool. After extracting data, we estimated the pooled sensitivity, specificity, likelihood ratios, and constructed summary receiver operating characteristics (SROC) curve. Results: Ten studies involving 367 malignant lesions and 178 benign lesions were considered eligible after full-text review. The pooled sensitivity and specificity were 0.95 (95% CI: 0.90-0.97) and 0.93 (95% CI: 0.85-0.97), respectively. Positive likelihood ratio and negative likelihood ratio were 12.8 (95% CI: 5.99-27.4) and 0.06 (95% CI: 0.03-0.11), respectively. The area under SROC curve was 0.98. Conclusions: Our meta-analysis indicates that DWI is a highly accurate diagnostic method in identification of colorectal cancer.

Keywords: MRI, diffusion magnetic resonance imaging, colorectal cancer, meta-analysis

Introduction

Colorectal cancer is the fourth leading cause of cancer-related death in both men and women all around the world [1]. It is most likely to occur among people over 40 years old and the prevalence of colorectal cancer is increasing steadily due to changes of lifestyles, especially in developing countries [2]. It is especially meaningful to detect lesions in the early stage, because early diagnosis of colorectal cancer associated with a much better prognosis for colorectal cancer therapy [3].

Various modalities, such as enteroscope biopsy, computed tomography (CT), and magnetic resonance imaging (MRI) have been used to detect colorectal lesions [4-6]. Although colonoscopy is an essential method for detecting colorectal cancer, it has many disadvantages. It is an invasive examination and need a long period of preparation before examination which may make patients receiving the examination uncomfortable. Besides, enteroscope biopsy also induces several complications, including bleeding, enterobrosis and tumor seeding. MRI seems to be a promising alternative for these investigations because there is no risk of radiation exposure, and no need for patient preparation as stated in ESGAR guidelines [7], and it is noninvasive. MRI has high resolution for soft tissues and can show the anatomic relationship clearly.

DWI was used to evaluate cerebral infarction in the acute stage initially. Recent years, DWI has emerged as a new diagnostic tool in different areas of radiology, particularly in the differentiation between malignant and benign diseases in the abdominal region. It is based on the principle of measuring the microscopic movements of water molecules in human tissues. Malignant tumors are composed of tumor cells that are randomly organized and form a dense mass. This structure characteristics prevents the free movement of water molecules and creates dif-
Detection of colorectal cancer

fusion restriction. Signal changes occur with the increased or limited diffusion movements of water molecules. Apparent diffusion coefficient can be calculated using this method and is quantitative measurement of diffusion ability [8]. Studies show that malignant tumors have statistically significantly lower ADC values than benign tumors. It has been reported that DWI has a high accuracy in diagnosis of prostate cancer [9], lung cancer [10], and bone marrow infiltration [11]. It is useful in preoperative staging and predicting response to neoadjuvant therapy in colorectal cancer [12, 13].

Some studies have shown that DWI could be an alternative imaging modality for initial detection of malignant tumors in the gastrointestinal region [14-18]. However, in the previous studies, the diagnostic performance of DWI in identification of colorectal cancer varied because of some factors such as field intensity, disease staging, lesion size, pathological type and so on. The sensitivity and specificity ranged from 82% to 100% and 65% to 100%, respectively. It still remains uncertain whether DWI is feasible for identification of colorectal cancer. Therefore, we performed this meta-analysis to systematically review the diagnostic value of DWI in identification of colorectal cancer.

Materials and methods

Search strategy

Our meta-analysis followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) recommendations. A comprehensive literature search of studies on human subjects was performed by two reviewers to identify articles about the diagnostic performance of DWI in detection of patients with colorectal cancer. PubMed and Embase were searched with the terms “diffusion weighted imaging [MeSH] or DWI [MeSH]” and “colorectal neoplasms [MeSH] or colorectal cancer or colorectal lesions or colon [MeSH] or rectum [MeSH]”. The latest search date was May 11, 2015, and there was no publication year limitation.

Eligibility criteria and study selection

Two investigators who were blinded to the journal, author, author’s affiliations and date of publication independently reviewed abstracts of all search results. The full-texts of eligible articles were subsequently retrieved for detailed reading. Studies were included if they met the following criteria: (a). DWI was conducted to diagnose colorectal cancer; (b). lesions were confirmed by pathologic findings (surgery specimens or endoscopic biopsy); (c). sufficient informations were available for calculation of the true-positive (TP), false-positive (FP), false-negative (FN), and true-negative (TN) values on a per-lesion basis; (d). the study population was no less than 10 patients. Studies were excluded if results were combination of different imaging modalities and data about a single modality could not be extracted.

Data collection and quality assessment

To perform this meta-analysis, the following data were extracted from each study: examination results, author, year of publication, country of origin, number and age of subjects, b values, techniques, and MRI field strength. Disagreements between the two reviewers were resolved by majority opinion after a third reviewer assessed all involved items. For each study, the published values for TP, FP, TN, FN, sensitivity, specificity, PLR, NLR, and diagnostic odds ratios were extracted. Results of TP, FP, TN, FN for the detection of colorectal cancer were extracted on a per-lesion basis. Eligible data were used to construct 2×2 contingency tables directly, with TP, FP, TN, FN results. In addition, a new method to extract the eligible data mentioned previously is applied in the diagnostic meta-analysis to improve the quantity of included studies and the amount of samples [19]. If both kinds of data were not available, we would contact the corresponding authors to ask for the detailed data. The study would be excluded without authors’ reply.

The methodological quality of the included studies was assessed by two independent observers using the QUADAS 2 tool for systematic reviews of diagnostic test accuracy [20].

Statistical analysis

Data extracted from each individual studies were used to construct 2×2 contingency tables, with the TP, FP, TN, FN values. The pooled summary sensitivity and specificity were estimated. PLR and NLR were derived as functions of these summary estimates. To explore heteroge-
neity, the $Q$ statistic and the inconsistency index ($I^2$) were obtained [21, 22]. If heterogeneity was not significant, the pooled analysis was performed using the fixed effects model. On the opposite condition, the random effects model was adopted [23]. The pooled sensitivity and specificity were calculated from individual studies with 95% CI. The SROC curve was constructed using the derived estimates of sensitivity and specificity. The area under the SROC curve (AUC) is used as evaluation of diagnostic ability of a test [24, 25]. Spearman correlation coefficient between the log (sensitivity) and log (1-specificity) was used to measure the threshold effect. Besides, threshold effect could be detected by plotting sensitivity and specificity on a ROC plane, the curve would have a typical pattern of “shoulder-arm” shape with the existing of it. If threshold effect existed, it is better to evaluate the accuracy of diagnostic test by fitting a SROC curve rather than pooling sensitivity and specificity. Subgroup analyses were also performed using single-factor meta-regression analysis. Publication bias was assessed visually by using Deeks’ funnel plot asymmetry tests, which is a scatter plot of the inverse of the square root of the effective sample size (1/ESS$^{1/2}$) versus the diagnostic log odds ratio. The non-zero slope coefficient was indicative of publication bias [26]. $P$ values of less than 0.05 were considered to be statistically significant. The statistical computations described above were performed using Stata/SE statistical software Version 12.0 (StataCorp LP, College Station, TX, USA).
Detection of colorectal cancer

Table 1. Clinical characteristics of included studies

<table>
<thead>
<tr>
<th>Study (first author)</th>
<th>Year</th>
<th>Country</th>
<th>Cancer patients</th>
<th>Non-cancer patients</th>
<th>Benign/Malignant</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hosonuma</td>
<td>2006</td>
<td>Japan</td>
<td>64.2/31-81</td>
<td>64.9/51-81</td>
<td>20/15</td>
<td>CR</td>
</tr>
<tr>
<td>Soyer</td>
<td>2010</td>
<td>France</td>
<td>69/43-84</td>
<td>65/30-81</td>
<td>31/31</td>
<td>R</td>
</tr>
<tr>
<td>Ichikawa</td>
<td>2006</td>
<td>Japan</td>
<td>59/33-69</td>
<td>-</td>
<td>33/15</td>
<td>CR</td>
</tr>
<tr>
<td>Rao</td>
<td>2008</td>
<td>China</td>
<td>60.9/21-86</td>
<td>-</td>
<td>20/45</td>
<td>R</td>
</tr>
<tr>
<td>Colosio</td>
<td>2013</td>
<td>France</td>
<td>64/-</td>
<td>14/8</td>
<td>22/27</td>
<td>CR</td>
</tr>
<tr>
<td>Lambregts</td>
<td>2010</td>
<td>Netherlands</td>
<td>68/35-87</td>
<td>64/22-81</td>
<td>23/19</td>
<td>R</td>
</tr>
<tr>
<td>Kilickesmez</td>
<td>2009</td>
<td>Turkey</td>
<td>57/-</td>
<td>45/-</td>
<td>39/14</td>
<td>CR</td>
</tr>
<tr>
<td>Nural</td>
<td>2013</td>
<td>Turkey</td>
<td>63/30-81</td>
<td>61/46-69</td>
<td>13/17</td>
<td>CR</td>
</tr>
<tr>
<td>Solak</td>
<td>2013</td>
<td>Turkey</td>
<td>57/31-77</td>
<td>-</td>
<td>15/26</td>
<td>CR</td>
</tr>
<tr>
<td>Avcu</td>
<td>2014</td>
<td>Turkey</td>
<td>-</td>
<td>-</td>
<td>31/27</td>
<td>C</td>
</tr>
</tbody>
</table>


Table 2. Imaging feature of each studies

<table>
<thead>
<tr>
<th>Study (first author)</th>
<th>Coil type</th>
<th>Company</th>
<th>Magnetic (T)</th>
<th>Methods</th>
<th>b value</th>
<th>ADC value (mm^2/s)</th>
<th>Patients</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hosonuma</td>
<td>body and spine matrix coil</td>
<td>SIEMENS AG</td>
<td>1.50</td>
<td>HSI; ADC</td>
<td>50; 800</td>
<td>1.19</td>
<td>1.37</td>
<td></td>
</tr>
<tr>
<td>Soyer</td>
<td>anterior torso phased-array coil</td>
<td>Siemens Healthcare</td>
<td>1.50</td>
<td>HSI; ADC</td>
<td>0; 500; 100</td>
<td>1.04</td>
<td>1.39</td>
<td></td>
</tr>
<tr>
<td>Ichikawa</td>
<td>Phased-array body coil</td>
<td>GE Healthcare</td>
<td>1.50</td>
<td>HSI</td>
<td>0; 1000</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Rao</td>
<td>Phased-array body coil</td>
<td>Siemens AG</td>
<td>1.50</td>
<td>HSI</td>
<td>0; 1000</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Colosio</td>
<td>abdominal phased-array coil</td>
<td>Siemens Healthcare</td>
<td>1.50</td>
<td>HSI</td>
<td>0; 500; 1000</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lambregts</td>
<td>Phased-array body coil</td>
<td>Philips Medical</td>
<td>1.50</td>
<td>HSI</td>
<td>0; 500; 1000</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kilickesmez</td>
<td>Phased-array body coil</td>
<td>Siemens Healthcare</td>
<td>1.50</td>
<td>ADC</td>
<td>0; 500; 1000</td>
<td>1.07</td>
<td>1.91</td>
<td></td>
</tr>
<tr>
<td>Nural</td>
<td>Phased-array body coil</td>
<td>Siemens Magnetom</td>
<td>1.50</td>
<td>HSI</td>
<td>0; 800</td>
<td>1.19</td>
<td>2.69</td>
<td></td>
</tr>
<tr>
<td>Solak</td>
<td>Phased-array body coil</td>
<td>Siemens Espree</td>
<td>1.50</td>
<td>HSI; ADC</td>
<td>0; 800</td>
<td>1.02</td>
<td>1.53</td>
<td></td>
</tr>
<tr>
<td>Avcu</td>
<td>Phased-array body coil</td>
<td>Siemens Magnetom</td>
<td>1.50</td>
<td>ADC</td>
<td>0; 800</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

HSI: Lesions were identified as malignant with the appearance of focal areas of high signal intensity. ADC: Lesions were identified as malignant with significant lower ADC value.

Results

The preliminary literature research included 203 potential article citations (Figure 1). One-hundred and thirty five of these studies were immediately excluded based on non-relevant, not being clinical trials, or publication in languages other than English or Chinese. Full-text of the remaining 68 articles were downloaded for detailed reading. 58 articles were excluded after the full-text reading. Finally, 10 published studies were included according to the inclusion and exclusion criteria [14-18, 27-31]. The study design characteristics were evaluated with the QUADAS 2 tool (Figure 2).

Of 203 studies, ten studies were included in the meta-analysis, involving 367 malignant lesions and 178 benign lesions. Among the ten studies, the anatomic location of lesions for six studies were rectum or colon, and three were rectum and one was colon. The mean age of patients in each individual studies ranged from 57 years old to 69 years old, and the median value was around 60 years old. The baseline characteristics of patients in these individual studies are summarized in Table 1. All images were obtained using the systems with 1.5 T magnetic field. Typical b-values for imaging were 0, 500, 800, and 1000 s/mm². There were two ways to identify a malignant lesion, one was by the existence of a local high signal intensity (HSI) area on the images of maximum intensity projection, and the other way was by the ADC value which was calculated from the signal intensity of region of interest on images with high b-value and low b-value. In three of the ten studies, malignant lesions were identified by both HSI and ADC value, and five were by HSI alone, and the remaining two were only by...
Detection of colorectal cancer

ADC value. The ADC value of malignant lesion ranged from 0.97 s/mm$^2$ to 1.19 s/mm$^2$, and benign lesion ranged from 1.37 s/mm$^2$ to 2.69 s/mm$^2$. Malignant lesions had a much lower ADC value. The imaging feature of each individual studies are shown in Table 2.

The sensitivity of DWI in identification of primary colorectal cancers ranged from 82% to 100% and specificity from 65% to 100%. The detailed data of sensitivity and specificity with 95% of CI for each individual study were given in Figure 3. To diagnose colorectal cancers, the pooled sensitivity and specificity were 0.95 (95% CI, 0.90-0.97) and 0.93 (95% CI: 0.85-0.97), respectively. Besides, the PLR and NLR ranged from 2.71 to 44.4 and 0.02 to 0.20, with pooled values of 12.8 (95% CI: 5.99, 27.4) and 0.06 (95% CI: 0.03, 0.11), respectively (Figure 4).

The test for threshold effect shows a Spearman correlation coefficient of 0.24 (P=0.51). The shape of ROC curve also doesn’t have “should-arm” pattern (Figure 5). These results indicate no significant threshold effect. The summary ROC curve (Figure 6) was fitted using data of each individual study, and AUC was 0.98, indicating a fine diagnostic accuracy.

The heterogeneity test of sensitivity and specificity reveals that Q equals to 14.6 (P=0.10, P=38.4%) and 25.9 (P=0.00, P=65.2%), respectively. And the heterogeneity of the positive and negative likelihood ratios are Q= 28.1 (P=0.00, P=46.8%) and 13.7 (P=0.13, P=34.5%), respectively. Thus except for sensitivity and negative likelihood ratio, other estimates exhibits some extent of heterogeneity. For homogeneous data, a fixed effects model was applied and the rest was analyzed using random effects model.

The nonsignificant slope of Deeks’ funnel plot asymmetry tests indicated that no significant bias was found (Bias =1.21, P=0.26), suggesting no major publication bias (Figure 7).

Figure 3. The sensitivity and specificity of DWI for the diagnosis of colorectal cancer.
Discussion

Colonoscopy plays a major role in the detection of colorectal cancer, and tissue samples for pathologic diagnosis could be obtained through endoscopic biopsy. However, the invasive nature of colonoscopy and need for patients preparation limited its clinical application [32]. CT is a widely used imaging method in abdominal and pelvic region. Its disadvantage is potentially harmful due to radiation and plain soft tissue resolution. It has been reported to have a low detection rate, and only good sensitivity for already highly advanced cases [33]. MRI seems to be a promising alternative for the investigations mentioned above because of the advantag-
es of no need for patient preparation, no risk of radiation, noninvasive and excellent soft tissue resolution. High-spatial-resolution T2WI MRI is increasingly used as an optimal method in detection of colorectal cancer. Signal intensity of T2WI MRI shows differences in the tumor, the mucosa and submucosal layers, the muscular layer, the perirectal fat, and the mesorectal fascia in rectum region. The tumor was usually identified as a region with signal intensity higher than the circular and longitudinal muscular layers and lower than the submucosa. Moreover, with the prolongation of examination time, more motion artifacts may be caused by excessive motion of the anterior abdominal wall and results in poor quality images [13].

DWI is a sensitive, functional imaging modality to detect the diffusion process of molecules, mainly water in tissue, in vivo and noninvasively [34]. It has been applied for early diagnosis of ischemic cerebral infarction over the last decade [35, 36]. Recently, it has been used in other clinical departments, especially in abdominal examination due to the development of ultra fast sequences such as echo planar imaging sequences (EPI) [37, 38]. It was reported that DWI could provide a desirable accuracy in the detection of prostate cancer [9], lung cancer [10], pancreatic cancer [39]. In colorectal cancer, it has been used in detection of lymphatic metastasis and predicting response to neoadjuvant therapy in colorectal cancer [11, 40]. However, the ability to differentiate benign and malignant colorectal diseases of DWI remains unclear and reports concerning this aspect is rare. Therefore, we firstly performed this meta-analysis.

Figure 6. Summary receiver operating characteristic (SROC) curve of DWI for differentiation of colorectal cancer.

Figure 7. Publication bias.
In this meta-analysis, we worked out the pooled sensitivity 95% (95% CI, 90 to 97%) and specificity 93% (95% CI, 85 to 97%) from the 10 included studies for the differentiation of benign and malignant colorectal lesions. In general, ranging between 0.50 and 1.00, AUC >0.80 indicates a good test. And the AUC we calculated by SROC was 0.97, significantly higher than the expected 0.80. Therefore, the result showed DWI had a great performance in identification malignant from benign colorectal lesions.

Some pooled estimates in this meta-analysis had heterogeneity. Threshold effect may be a major cause of it, however, the Spearman’s correlation test shows no significant threshold effect, and the shape of ROC is also not typical “shoulder arm” pattern. We considered the possibility of publication bias as the origin of heterogeneity. However, Deeks’ funnel tests indicated no publication bias (P=0.26). To further investigate the source of heterogeneity, we performed subgroup analysis. The result showed there were no considerable improvement of homogeneity in subgroups divided according to country, b-value, imaging techniques and so on. In the subgroup excluding the study of Hosonuma whose spectrum of patients wasn’t very proper shows a great improvement in homogeneity [27]. Hosonuma’s study has the disadvantage of having patients with colon cancer as control group when evaluate the diagnostic value of DWI in identification of rectal cancer. The subgroup analysis shows that sensitivity and specificity were 0.94 (Q=10.5, P=0.12, I²=37.5%) and 0.93 (Q=10.7, P=0.22, I²=25.2%), respectively. Therefore, heterogeneity decreases a lot than initial pooled estimates. In the subgroup analysis, although some factors such as histological type, quality of imaging system, professional ability of radiologist doctor may still cause slight heterogeneity, our meta-analysis has a relatively desirable homogeneity.

In summary, all currently available evidence indicates that DWI is an accurate, noninvasive, and non-radiative imaging technique for distinguishing malignant from benign colorectal lesions. In the future, however, large-scale randomized control trials are necessary to further validate the clinical value of DWI and establish standards of measurement, analysis, and cut-off values of diagnosis.

Disclosure of conflict of interest
None.

Address correspondence to: Drs. Xuelei Ma and Feng Wang, Department of Oncology, West China Hospital, No. 37, Guoxue Alley, Chengdu 610041, PR China. Tel: +86 28 85475576; E-mail: maxlhx@163.com (XLM); fengwanghx@gmail.com (FW)

References
Detection of colorectal cancer


