The diagnostic accuracy of B-mode ultrasound in detecting meniscal tears: a systematic review and pooled meta-analysis

Fajin Dong1*, Lei Zhang1*, Shuxia Wang2, Duo Dong3, Jinfeng Xu1 Huaiyu Wu1, Yan Liu4, Min Li5

*the authors shared the first authorship
1Department of Ultrasonography, Second Clinical College of Jinan University, Shenzhen People’s Hospital, Shenzhen, 2Department of Ultrasonography, Zhucheng People’s Hospital, Shandong, 3Domestic Clinical Application Department(UIS) of Shenzhen Mindary Bio-Medical Electronics CO., LTD, 4Department of Cardiology, Qilu Hospital, Shandong University, Shandong, 5Department of Ultrasonography, Second Clinical College of Jinan University, Shenzhen People’s Hospital, Shenzhen, China.

Abstract

Aim: To evaluate the diagnostic accuracy of meniscal tears using B-mode ultrasound and high-frequency linear probe by conducting a systematic review and pooled meta-analysis. Material and methods: The Cochrane library, Embase, and Pubmed were searched for relevant studies up to 29 July 2017. The arthroscopy was used as the reference standard. The results were estimated by pooled sensitivity, specificity, diagnostic odds ratio, likelihood ratio, and the area under the summary receiver operating characteristic (SROC). Results: Seven prospective studies met the selection criteria, comprising 321 meniscal tears from 472 patients. The pooled sensitivity, specificity, positive likelihood ratio, negative likelihood ratio, diagnostic odds ratio, and area under the SROC curve were 88.80% (95%CI: 82.83-92.87), 84.66% (95%CI: 75.89-90.64), 5.79(95%CI: 3.66-9.15), 0.13 (95%CI: 0.09-0.20), and 43.74 (95%CI: 24.01-79.68), respectively. The area under the SROC curve was 93% (95%CI: 91-95). Conclusions: This meta-analysis indicates that 2-dimensional ultrasound is useful, and could be routinely used for estimating meniscal injuries in the human knee joint. Keywords: B-mode ultrasound; meniscal tear; meta-analysis.

Introduction

The menisci of the knee are two pads of fibrocartilaginous tissue which serve to disperse friction in the knee joint between the tibia and femur. Meniscal tears of knee are the most common lesions in sports activities due to sudden changes in direction of running, or jumping. They often involve the medial meniscus, as it is less mobile than the lateral meniscus [1].

Clinical diagnosis of meniscal tears can be difficult even for the most experienced orthopaedic surgeons, which is most often diagnosed based on history, clinical symptoms, magnetic resonance imaging (MRI), or arthroscopy. MRI can be considered as the non-invasive “gold standard” for the diagnosis of meniscal tears [2], but it is expensive and needs long examination times. Arthroscopy is considered as the gold standard [3], but it is invasive, expensive, and requires a surgery admission.

Diagnostic ultrasound (US) of the knee can identify abnormalities in the menisci. US examination, especially B-mode US with liner probes, has been tried for the assessment of meniscus injuries with variable results. It is simple, convenient and an inexpensive and non-invasive method [4-11]. However, its accuracy remains controversial [12].

Two meta-analysis concerning the diagnosis of meniscal tears using US [13,14] have been published in the recent years, but in these papers no difference between the types of used transducers (high-frequency linear probe, low-frequency convex array probe, 3-dimensional
probe) was made. For this reason we performed a systematic review and meta-analysis to investigate the diagnostic performance of meniscal tears by using B-mode US and high-frequency linear probe.

**Material and methods**

**Search strategy**

The Cochrane library, Embase, and Pubmed were searched for relevant studies up to 29 July 2017. Mesh terms and free words were used, Meniscus Injuries or meniscus tears or meniscal tears or ultrasound etc. The search strategy was (((((((Meniscus Injuries[MeSH Terms]) OR meniscus tear) OR meniscus tears) OR meniscal tear) OR meniscal tears)) AND (((Ultrasonography[MeSH Terms]) OR Ultrasonics[MeSH Terms]) OR ultrasound)) that we used in Pubmed, Embase and Cochrane library. The literature search was performed without language restrictions. Moreover, the reference lists of the retrieved systematic and narrative reviews were also manually searched to identify additional relevant studies.

Only studies evaluating the diagnostic accuracy of meniscal tears by using B-mode US (and with high-frequency linear probe) were included. The arthroscopy was used as the reference standard. The results were estimated by pooled sensitivity, specificity, diagnostic odds ratio, likelihood ratio, and the area under the summary receiver operating characteristic (SROC).

**Study selection**

The inclusion criteria were: prospective or retrospective cohort design; human studies; study population at least 20 patients; evaluation of the meniscal tears by using B-mode US and high-frequency linear probe; use of arthroscopy as the reference standard; providing the true-positive (TP), true-negative (TN), false-positive (FP), and false-negative (FN) number at per study. If the studies did not provide data to directly construct 2×2 contingency tables, we calculated from the reported sensitivity, specificity, negative predictive value, and positive predictive value. Studies such as reviews, meta-analysis, letters, abstracts, case reports, or editorials were excluded. Two reviewers (Y. L. and J.F. X.) with similar level of experience and expertise independently screened the eligible studies. Disagreements between two reviewers were resolved by discussion or consensus from a third reviewer (F.J. D.), who rechecked the search results and the assessment process.

**Data extraction and quality assessment**

The following informations were independently extracted by two authors (L. Z. and Z.M. D.) using a standardized form: author, year, country, design, number of patients, age range (y), ultrasound probe, reference standard, TP, FP, FN, TN, sensitivity, and specificity.

The methodological quality of eligible studies was assessed by using the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool [15]. The QUADAS-2 tool consists of 4 key domains and 3 applicability concerns. The key domains concern the patient selection, index test, reference standard, flow and timing. Applicability concerns are structured in a way similar to risk bias sections but not to include the flow and timing. For each item of the QUADAS-2 tool, if a study is judged as “low” on all domains relating to bias or applicability, then it is appropriate to have an overall judgment of “low risk of bias” or “low concern regarding applicability” for that study. If a study is judged “high” or “unclear” in one or more domains, then it may be judged “at risk of bias” or as having “concerns regarding applicability.”

**Statistical analysis**

All the analysis was performed by using the Midas module of Stata statistical software, version 14.0 (Stata Corp, College Station, TX, USA), RevMan software, version 5.3.5 (The Nordic Cochrane Center, The Cochrane Collaboration, Copenhagen, Denmark). Stata 14.0 with the bivariate mixed-effects regression model developed by van Houwelingen, modified for the synthesis of diagnostic test data [16,17], was used to pool statistics indexes and draw statistical graphs. The pooled sensitivity (P-SEN) and specificity (P-SPE), pooled positive likelihood ratio (PLR+), negative likelihood ratio (PLR-), and diagnostic odds ratio (DOR) with corresponding 95% confidence intervals (CI), area under the summary receiver operating characteristic (SROC) curve were used to examine the diagnostic accuracy. RevMan 5.3.5 was used to assess the methodological quality of the eligible studies.

The PLR+ and PLR- were calculated based on the P-SEN and P-SPE by assuming prior probabilities of 20%. The Inconsistency index (I²) and the Cochrane Q statistic value were used to estimate the heterogeneity across the included studies. If the Cochran Q statistic value was p>0.1 or I²<50%, we used a fixed-effect model; otherwise, a random effects model was selected [18].

**Publication bias**

Deek’s funnel plot asymmetry test was used to test the potential publication bias, with p<0.10 for the slope coefficient indicating significant asymmetry [19]. It was conducted by a regression of diagnostic log odds ratio (lnDOR) versus the inverse of the square root (1/sqrt) of the effective sample size (1/ESS1/2) and weighted by the effective sample size. A p-value <0.10 for the slope coefficient indicated significant asymmetry. Furthermore, groups were divided into subgroups based on the heterogeneity between studies.
Results

Literature searches

The initial search yielded 1771 studies. After screening the title or abstract and excluding the duplication, 7 full-text studies were assessed for eligibility and ultimately included in this meta-analysis (fig 1).

Study characteristics

Seven prospective diagnostic cohort studies met the selection criteria [5-11], involving 321 meniscal tears from 472 patients. Table I summarizes the detailed characteristics of the included studies.

Methodology quality assessment

According to the methodological assessment of the QUADAS-2 checklist, the quality of the included studies was judged to be high and all the quality assessment items of the included studies had a low risk of bias (fig 2).

Data synthesis and analysis

As shown in figure 3, significant heterogeneity in P-SEN (I²=53.14%, Q=12.80, p=0.05>0.01) and P-SPE (I²=35.92%, Q=9.36, p=0.15>0.01) was detected, so we used a fixed-effect model.

The P-SEN, P-SPE, PLR+, PLR-, DOR, and area under the SROC curve were 88.80% (95%CI: 82.83–92.87), 84.66% (95%CI: 75.89–90.64), 5.79 (95%CI: 3.66–9.15), 0.13 (95%CI: 0.09–0.20), 43.74 (95%CI:24.01–79.68), and 93% (95% CI: 91–95) respectively (fig 4-7).

Table I. Characteristics and diagnostic performance of included studies

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Country</th>
<th>Design</th>
<th>No. of Patients</th>
<th>Ultrasound Probe</th>
<th>Reference Standard</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
<th>TN</th>
<th>Sen</th>
<th>Spe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Khan 2006</td>
<td>Saudi Arabia</td>
<td>P</td>
<td>60</td>
<td>7.5 MHz LP</td>
<td>Arthroscopy</td>
<td>37</td>
<td>2</td>
<td>3</td>
<td>18</td>
<td>0.925</td>
<td>0.900</td>
</tr>
<tr>
<td>Sandhu 2007</td>
<td>India</td>
<td>P</td>
<td>51</td>
<td>7.5 MHz LP</td>
<td>Arthroscopy</td>
<td>40</td>
<td>3</td>
<td>2</td>
<td>6</td>
<td>0.952</td>
<td>0.667</td>
</tr>
<tr>
<td>Shetty 2008</td>
<td>England</td>
<td>P</td>
<td>35</td>
<td>5–13 MHz LP</td>
<td>Arthroscopy</td>
<td>19</td>
<td>4</td>
<td>3</td>
<td>9</td>
<td>0.864</td>
<td>0.692</td>
</tr>
<tr>
<td>Alizadeh 2013</td>
<td>Iran</td>
<td>P</td>
<td>74</td>
<td>14-MHz LP</td>
<td>Arthroscopy</td>
<td>53</td>
<td>3</td>
<td>5</td>
<td>13</td>
<td>0.914</td>
<td>0.813</td>
</tr>
<tr>
<td>Timotijevic 2014</td>
<td>Serbia</td>
<td>P</td>
<td>107</td>
<td>5–7.5 MHz LP</td>
<td>Arthroscopy</td>
<td>41</td>
<td>6</td>
<td>13</td>
<td>47</td>
<td>0.759</td>
<td>0.887</td>
</tr>
<tr>
<td>Lan 2006</td>
<td>China</td>
<td>P</td>
<td>74</td>
<td>7.5–10 MHz LP</td>
<td>Arthroscopy</td>
<td>38</td>
<td>1</td>
<td>8</td>
<td>27</td>
<td>0.826</td>
<td>0.964</td>
</tr>
<tr>
<td>Cook 2014</td>
<td>America</td>
<td>P</td>
<td>71</td>
<td>10-14 MHz LP</td>
<td>Arthroscopy</td>
<td>54</td>
<td>2</td>
<td>5</td>
<td>10</td>
<td>0.912</td>
<td>0.842</td>
</tr>
</tbody>
</table>

P-Prospective; LP – Linear Probe; TP – true-positive; TN – true-negative; FP – false-positive; FN – false-negative; Sen – sensitivity; Spe – specificity
Fig 3. Forest plot to demonstrate study-specific on the right y-axis for sensitivity and specificity of B-mode ultrasound with a high-frequency linear probe in evaluating meniscal tears.

Fig 4. Forest plot to demonstrate study-specific on the right y-axis for a negative and positive likelihood ratio of B-mode US with a high-frequency linear probe in evaluating meniscal tears.

Fig 5. Forest plot to demonstrate study-specific on the right y-axis for diagnostic odds ratio of B-mode US with a high-frequency linear probe in evaluating meniscal tears.

Fig 6. Hierarchical summary receiver operating characteristic curve (HSROC) with summary point, summary estimates, 95% confidence region and 95% prediction region for all included studies of B-mode US with a high-frequency linear probe in evaluating meniscal tears (n = 7).

Fig 7. Fagan’s Nomogram with pretest probability at 20%.
Publication bias

The Deeks’ funnel plot for testing publication bias showed that the studies were distributed symmetrically with a p-value of 0.99, indicating no clear evidence of publication bias (fig 8). Regression analysis of lnDOR against 1/Effective Sample Size\(^{1/2}\) showed no obvious small-study bias in our meta-analysis.

Discussions

Knee injuries are more frequent in young people, mainly males, probably due to the higher proportion of men compared to women who engage in sports and sporting activities. Although MRI is currently the most reliable noninvasive diagnostic method of choice in evaluation of meniscal tears, it is an expensive imaging method and patients generally have a long waiting period for schedule.

Menisci tears have many different shapes (vertical, horizontal, radial, oblique, complex) and sizes. The middle and posterior parts of the menisci can be easily damaged. In recent years, the image quality of US technique, especially high resolution US, has dramatically improved. Also the method is safe, convenient, inexpensive, easily accessible, and can be performed rapidly. But the effectiveness of diagnosis in meniscus tears is still controversial. Azzoni et al [20] published a retrospective study in 321 patients using a low-frequency transducer and concluded that ultrasound was neither sensitive (sensibility 60%) nor specific (specificity 21%) for diagnosing meniscal tears. The same low sensitivity (32%) and specificity (59%) was obtained by Bruce et al in 56 patients [21]. But other studies had reported promising results [12,22,23]. Our study demonstrated that B-mode US with high-frequency linear probe had high P-SEN, P-SPE, PLR+, PLR-, DOR, and area under the SROC.

In our meta-analysis, moderate to significant heterogeneity was present in the summary results. We postulated that several factors may contribute to the good heterogeneity. First, the US resolution has been improved; second, the resolution of the linear array probe is higher compared with the convex array probe; third, the musculoskeletal US is becoming more and more popular and the technique is more and more skillful. For these reasons, more patients with meniscal tear are able to be detected.

To our knowledge, this is the first meta-analysis exploring the diagnostic accuracy of high-resolution US in the detection of meniscal tears. The quality of the selected studies was higher, as all had QUADAS full scores. All the selected studies were prospectively designed, and high-resolution US linear probes were used, which reduced the potential bias and guaranteed the robustness of the results.

Limitations

Nevertheless, this meta-analysis does have some limitations. Although 7 studies were included, the number of patients (321 meniscal tears) is still relatively small so we combined the lesions of medial and lateral meniscal tears (medial meniscus injury was easier diagnosed than the lateral). Some of the papers included in our study analyzed individually the medial and lateral meniscus and our global analysis may introduce bias. As we excluded from the analysis the papers in which the transducer was 3-dimensional or convex array, and the retrospective studies, the real value of US in the diagnosis of meniscal tears was not entirely established. Another possibility for bias is related to the different quality of the US equipment used in each study. In order to further validate the diagnostic accuracy of US in meniscal injury, a larger prospective study, with uniform US equipments, is warranted.

This meta-analysis demonstrates that the diagnostic accuracy of US for diagnosing meniscal injury is acceptable, with a relatively high specificity and sensitivity. Therefore, besides other techniques, such as MRI or arthroscopy, US may be a complementary method to enhance the sensitivity of the evaluation.

Conclusions

In summary, our meta-analysis suggests that B-mode US with a high-frequency linear probe is an acceptable imaging technique in evaluating meniscal tears. Due to the high sensitivity, specificity, and diagnostic accuracy,
US can be considered as a useful method to detect meniscal tears.

Acknowledgements: This project was supported by the Clinical Research Project of Shenzhen People’s Hospital (No. SYLY201702).

Conflicts of interest: none

References