Ultrasound examination in juvenile idiopathic arthritis is better than clinical examination for identification of intraarticular disease

Hans Erik Nielsen1, Charlotte Strandberg2, Steen Andersen3, Christina Kinnander2 & Gunna Erichsen1

ABSTRACT
INTRODUCTION: The diagnosis of juvenile idiopathic arthritis (JIA) is formally based on clinical examination, but ultrasound (US) examination is used increasingly. Our purpose was to compare US and clinical examination in the assessment of synovitis in JIA.

MATERIAL AND METHODS: This study was retrospective and included 62 consecutive patients with newly diagnosed JIA admitted to the Department of Paediatrics at Gentofte Hospital, Denmark from 2003 to 2010. The included patients were examined clinically and by US at their first visit. All peripheral joints were examined clinically, 24% of these joints were examined by US. The development of new, clinically arthritic joints was followed during the next six months.

RESULTS: The mean sensitivity of clinical examination was 48% with a clear hierarchy among joints, knees having the highest sensitivity, small joints of hands and feet the lowest. On average, 0.4 joints per child, which were arthritic by clinical examination, were negative by US. Inversely, US detected 1.3 more arthritic joints than clinical examination did per child. The latter is a minimum estimate since only 24% of the joints were examined by US. Subclinically arthritic joints had a 29% probability of developing clinical arthritis within the first six months following the initial examination.

CONCLUSION: Although there is no formal validation of US examination in children suspected for JIA, we recommend that it is used routinely and performed by a highly experienced US operator.

FUNDING: not relevant.
TRIAL REGISTRATION: not relevant.

Juvenile idiopathic arthritis (JIA) is a heterogeneous group of chronic arthritis conditions affecting children under the age of 16 years. The diagnosis of JIA and its classification into oligoand polyarticular disease is, for historical reasons, based on clinical examination [1], because ultrasound (US) and magnetic resonance imaging (MRI) have only become available in recent years. MRI is more sensitive than US for detection of early intraarticular inflammation and evaluation of late manifestations of the disease [2-4]. MRI is not feasible for routine studies in children. US is considered accurate for the identification of intraarticular disease. It is possible to examine multiple joints with US and to detect even subclinical synovitis [5-10].

In JIA, two studies of multiple joints [5, 6] and several studies of selected joints [8, 9, 11-13] have suggested that clinical examination is not reliable if US examination is accepted as the standard. Similar findings have been obtained by comparison of clinical evaluation with MRI [2, 3]. The large inter-observer variation associated with clinical examination [12, 13] also suggests that it is unreliable.

The validation of US diagnosis in JIA was reviewed by Collado et al [8]. They concluded that a formal validation is lacking, but they consider US to be useful and superior to clinical examination.

The purpose of this study was to perform a general comparison of clinical and US evaluation in the assessment of articular inflammation in JIA at the first examination in a routine, clinical setting. Furthermore, the aim was to determine the precision of clinical evaluation in different joints. Finally, the aim was to examine if demonstration by US of subclinical synovitis in a joint predicts the development of clinical arthritis in that joint within the following six months.

MATERIAL AND METHODS
Patients
All consecutive patients who satisfied the clinical criteria for JIA of the International League of Associations for Rheumatology (ILAR) [1] and who were admitted during the period from January 2003 to April 2010 to the Department of Paediatrics at Gentofte Hospital, Denmark, were included. The results of the first clinical and US examination were compared. US examination was performed in all cases of suspected rheumatic disease.

This study was retrospective. Patients were identified from diagnostic registers of the paediatric department and the radiological information system. The examiners were not blinded.

All patients were examined by one of two experienced rheumatologists and by one of two radiologists specialized in paediatric musculoskeletal ultrasound. For
the calculations, arthritis was not graded, but classified as either positive or negative in the clinical as well as the US examination.

Inclusion criteria:

- Newly diagnosed JIA (diagnosis normally later than at the first examination).

Exclusion criteria:

- More than 14 days between clinical and US examination.
- Systemic JIA (joint manifestations often occur late).
- Preceding treatment with disease-modifying antirheumatic drugs (DMARDs) or intraarticular steroid.

Clinical examination
The affected joints, as noted by the patient and/or parents, were registered. All peripheral joints were examined and classified according to the ILAR criteria [1] as arthritic if there was intraarticular swelling or non-arthritic if no swelling was present there was pain/tenderness on motion and limitation of motion. All children were examined with the same unchanged technique during the study period.

The development of new, arthritic joints, defined by clinical examination, was registered during the six months following the initial examination. Intervals between clinical controls were six weeks or less.

Ultrasound examination
The US examination was performed with a high-end US scanner (General Electric Logic 9 or Esaote XV Gold) using a high-frequency linear probe 9-15 MHz. The Doppler was adjusted for lowest flow.

Either colour or power Doppler was used to evaluate synovial blood flow, and the method with the highest sensitivity for each scanner was selected. Hyperaemia was classified as positive or negative.

Images of all examined joints were stored in the picture archiving and communication system (PACS) and described by a uniform template.

In case of subjective symptoms or clinical signs from the upper extremities, symptomatic joints, elbows and wrists were normally examined. Similarly, in the lower extremity, symptomatic joints, the hips, knees and ankles were normally examined.

During the period of study, the routine scanning protocol was expanded because we discovered that more joints than suspected clinically were involved in several children. During the last four years of the study, the routine scanning protocol therefore included all the joints of the lower extremity, the wrists and elbows. Similarly, during the first two years of study, the sensitivity to detect synovial blood flow with Doppler was low, and for that reason it was not used routinely to evaluate hyperaemia. However, after upgrading with new software in 2005, all joints with hypoechoic synovial hypertrophy were examined with Doppler.

The technique and the diagnostic criteria were otherwise unchanged during the study period. The US data were extracted from the stored descriptions and when necessary the stored images were reviewed. It was a routine to perform a follow-up US examination of the same joints as those that were examined at the first US examination in addition to the clinical control. Furthermore, new, clinically arthritic joints were examined by US.

The definitions of US pathology used were those of the OMERACT 7 group [7], which are also applied for children [9]. Joint effusion is defined as abnormal anechoic or hypoechoic (relative to subdermal fat) intraarticular material that is displaceable, compressible and without Doppler activity (Figure 1 A).
Synovial hypertrophy is defined as abnormal hypoechoic (relative to subdermal fat), but sometimes isoechoic or hyperechoic, intraarticular tissue that is nondisplaceable and poorly compressible (Figure 1B). Doppler activity is a marker for the degree of inflammation.

Tenosynovitis is defined as hypoechoic or anechoic, thickened tissue with or without fluid in the tendon sheath, which is seen in two perpendicular planes and which may exhibit Doppler signal [7] (Figure 2).

**Trial registration:** not relevant.

**RESULTS**

A total of 90 children with JIA were admitted during the study period. In all, 28 of these were excluded, 25 due to a time difference of more than two weeks between the clinical and the US examination, and three due to systemic JIA. This left 62 children for study. The basic data are shown in Table 1. A total of 4,464 joints were examined clinically, and 1,064 by US; consequently, a mean of 17 joints per child were examined by US. The joints of the carpus and the tarsus were only counted as one joint.

US detected more arthritic joints than the clinical examination; on average 1.3 more joints per child; these were considered false negative joints. The clinical examination also produced false positives; on average 0.4 joints per child. The frequency of false positive joints was highest for the wrists (5%) and the ankle (6%).

The classification by US and clinical examination differed in six children. In five the clinical diagnosis was oligoarthritis, while US showed polyarticular disease. In one, the clinical diagnosis was polyarthritis, while US showed oligoarticular disease.

Table 2 shows that there are large differences between individual joints with regard to sensitivity and specificity of the clinical examination. The pooled sensitivity for all joints is 48%, highest for the knee (69%) and lowest for the carpal and tarsal joints, and the small joints of fingers and toes. The sensitivity is only 36% for the hip joints.

The specificity varied between 92% and 100%. The differences of specificity between individual joints were not statistically significant.

We could not estimate the interobserver variation directly, but we repeated the calculations of sensitivity and specificity for the four different combinations of clinical and US examiner. There were no significant differences among them.

We made additional subanalyses comparing the following: The first and second study period, before and after introduction of Doppler. Those with clinically monoarticular disease compared with those with more than one arthritic joint. Those with and without elevated sedimentation reaction. There were no significant differences in sensitivity in any of these comparisons, but the power was limited due to small numbers.

Tenosynovitis was demonstrated by US around 16 ankles, in 11 without intraarticular synovitis. For comparison, there were 13 ankles with intraarticular synovitis in five of which tenosynovitis was also present. Tenosynovitis was also seen in the wrists, although rarer – two joints with isolated tenosynovitis were observed as well as 18 joints with intraarticular synovitis; seven of the latter with simultaneous tenosynovitis.

At the first examination, 35 of the 62 children had clinically silent joints, i.e. clinically nonarthritic joints in which synovitis was demonstrated by US examination. A total of 80 clinically silent joints were found. Of these, 23 joints developed clinical arthritis within the next six months, a frequency of 29%. For comparison, within six months, these 35 children developed clinical arthritis in nine of the 847 joints which were normal at the first clinical and US examination, a frequency of 1%.

**DISCUSSION**

Our study is presently the largest published on this subject. It is one of few studies [10] with follow-up, and the only large study that routinely performed US in relation to the first clinical visit. It reflects the findings of experienced examiners in a routine clinical setting, but the study design has some limitations: it is retrospective, the examiners were not blinded, we did not include normal controls and there was no fixed protocol for the US examination.

The fact that only 24% of the joints were examined by US does not affect the estimate of 0.4 false positive joints per child, since all clinically active joints were examined by US. However, it may lead to too low esti-

### Table 1

<table>
<thead>
<tr>
<th>Basic patient data.</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients included, total, n</td>
<td>62</td>
</tr>
<tr>
<td>Boys/girls, n</td>
<td>22/40</td>
</tr>
<tr>
<td>Age, years, median (range)</td>
<td>9 (1-15)</td>
</tr>
<tr>
<td>Duration of symptoms prior to inclusion, months, median (range)</td>
<td>4 (0-120)</td>
</tr>
<tr>
<td>Numeric time interval between US and clinical exam, days, median (range)</td>
<td>0 (0-14)</td>
</tr>
<tr>
<td>US exam prior to clinical exam, %</td>
<td>29</td>
</tr>
<tr>
<td>ANA-positive/total, n</td>
<td>14/62</td>
</tr>
<tr>
<td>HLA-B27-positive/examined, n</td>
<td>8/47</td>
</tr>
<tr>
<td>Oligo-/polyarticular by initial clinical exam, n</td>
<td>58/4</td>
</tr>
<tr>
<td>Joints examined clinically, total/by US, n⁴</td>
<td></td>
</tr>
</tbody>
</table>

ANA = antinuclear antibodies; HLA = human leukocyte antigen; US = ultrasound.

a) Main reason for large difference between numbers: small joints of fingers and toes were not routinely scanned.

b) Carpal and tarsal joints only counted as 1 joint.
The sensitivity reported when multiple joints were examined in those studies was small. The sensitivity is reported [6, 10, 14]. However, the number of joints examined in those studies was small. The sensitivity is lower for all other joints; notably the hip is difficult to examine clinically, the sensitivity of hip examination being 36% in our study. Similar comparisons of US and clinical examination of the hip, allowing for estimation of sensitivity, have not been reported, but the sensitivity of clinical examination of the hips was only 26% compared with MRI examination [3].

There may be several explanations for the low sensitivity of the clinical examination: The agreement between two different clinical examiners is limited [14-16]; it is better, but far from perfect among two US examiners [5, 12, 13]. Clinically, it is not possible to distinguish between synovitis and joint effusion, tenosynovitis and subcutaneous oedema.

Our finding that synovitis demonstrated by US in clinically silent joints predicts development of clinical synovitis, strongly supports the biological significance of the US findings. There are several reports of subclinical arthritis demonstrated by US [2, 5, 6, 8, 11] and by MRI [2], two of which [2, 10] included follow-up and reported a risk of development of clinical arthritis similar to ours (36 and 20%). The relatively low number of US-positive joints that develop clinical arthritis does not necessarily mean that the US results are false positives; the joints may have responded to treatment.

The critical question in this and other similar studies is the validity of US diagnosis of synovitis, which has not been established; although US is widely used and considered to be valuable [8]. In adults there is some direct evidence of the validity from comparison of US and synovial histopathological findings in rheumatoid arthritis [17-19]. Similar studies have not been – and will hardly be – performed in children.

US is inexpensive, easy to perform and to repeat. Furthermore, US is nonionizing, useful in all age groups and yields a more precise classification than physical examination. The disadvantage is that US examination is very operator-dependent – as is clinical examination.

In young children, a large part of the joint consists of cartilage which can be difficult to distinguish from synovitis. For that reason synovitis may be overestimated by US which may lead to false positive joints unless the operator has routine in examination of children.

One study in children and several studies in adults have made quantitative comparisons of US and MRI findings in arthritis; a few of these also included clinical examination in the comparisons [4, 8].

Generally, the correlation between US and MRI is good and better than the correlation between MRI and clinical findings.

In conclusion, our main findings are: That clinical examination is unreliable for demonstration of synovitis in JIA when US results are taken as the standard, that the sensitivity of clinical examination varies among groups of joints, and that demonstration of synovitis by US in a clinically silent joint predicts development of clinical arthritis in that joint.

The evidence from our study and from the literature supports routine use of US in JIA, and we suggest that US performed by an experienced operator and according to a standardized protocol should be used routinely as a supplement to clinical examination in JIA. However, further prospective studies comparing US and MRI findings in JIA are required to demonstrate the validity of US.

### Table 2

Sensitivity and specificity of the clinical exam with US findings as the standard.

<table>
<thead>
<tr>
<th>Joints examined both clinically and by US, n^a</th>
<th>Sensitivity, % (95% CI)</th>
<th>Specificity, % (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shoulder</td>
<td>10</td>
<td>0 (0-84)</td>
</tr>
<tr>
<td>Elbow</td>
<td>64</td>
<td>55 (23-83)</td>
</tr>
<tr>
<td>Wrist</td>
<td>60</td>
<td>50 (26-74)</td>
</tr>
<tr>
<td>Carpal joints</td>
<td>40 hands</td>
<td>0 (0-84)</td>
</tr>
<tr>
<td>MCP and finger joints</td>
<td>312</td>
<td>30 (7-65)</td>
</tr>
<tr>
<td>Hip</td>
<td>112</td>
<td>36 (21-53)</td>
</tr>
<tr>
<td>Knee</td>
<td>118</td>
<td>69 (55-81)</td>
</tr>
<tr>
<td>Ankle</td>
<td>106</td>
<td>54 (25-81)</td>
</tr>
<tr>
<td>Tarsal joints</td>
<td>64 feet</td>
<td>11 (0-48)</td>
</tr>
<tr>
<td>MTP and toe joints</td>
<td>184</td>
<td>9 (0-41)</td>
</tr>
<tr>
<td>Average, all joints</td>
<td>1,064</td>
<td>48</td>
</tr>
</tbody>
</table>

CI = confidence interval; MCP = metacarpo-phalangeal; MTP = metatarso-phalangeal; US = ultrasound.

a)Reason for the different numbers: in some patients only upper or lower extremities were US examined.

b) A common CI is given as there were no significant differences between the specificities of different joints.

c) 95% CI for sensitivity and specificity of the clinical exam were calculated from the binomial distribution.
CORRESPONDENCE: Hans Erik Nielsen, Hostrup Have S8, 1954 Frederiksberg C, Denmark. E-mail: grethehans@mail.tele.dk

ACCEPTED: 14 May 2013

CONFLICTS OF INTEREST: none. Disclosure forms provided by the authors are available with the full text of this article at www.danmedj.dk.

LITERATURE