A new magnetic resonance imaging scoring method for assessment of haemophilic arthropathy.

Lundin, Björn; Pettersson, Holger; Ljung, Rolf

Published in:
Haemophilia

DOI:
10.1111/j.1365-2516.2004.00902.x

Published: 2004-01-01

Citation for published version (APA):
A new magnetic resonance imaging scoring method for assessment of haemophilic arthropathy

B. LUNDIN,* H. PETTERSSON* and R. LJUNG†
*Department of Radiology, University Hospital of Lund, Lund; and †Departments of Paediatrics and Coagulation Disorders, University Hospital of Malmö, Malmö, Sweden

Summary. In a European multicentre study, 39 ankles in 28 haemophilic boys were investigated by magnetic resonance imaging (MRI). A new MRI score was developed in the format A(e:s:h) for evaluating haemophilic arthropathy. This scheme provides high resolution and allows separation of different pathological components. The factor A is calculated as the sum of scores for subchondral cysts (maximum value 6), irregularity/erosion of subchondral cortex (maximum 4) and chondral destruction (maximum 6); e, s and h, respectively, represent effusion/haemarthrosis, synovial hypertrophy and haemosiderin deposition, and they are separately evaluated on a scale of 0–4. Working independently, two radiologists scored the 39 ankles twice using both this new ‘European’ scoring method and a previously published ‘Denver’ scoring scheme. Final classification was achieved by consensus. The reproducibility of the readings was assessed, and for both scoring methods the results indicated good or moderate intraobserver agreement, and good, moderate or fair interobserver agreement. These findings suggest that MRI can be useful for semiquantitative evaluation of haemophilic arthropathy, providing the examination is performed according to an appropriate protocol, and the images are evaluated by specially trained radiologists.

Keywords: factor IX, factor VIII, haemophilia, haemophilic arthropathy, MRI

Introduction

Haemophilic arthropathy is caused by joint bleeding, and the rate of progression of this disease is determined mainly by the number of haemarthroses [1–4]. The therapeutic action of greatest importance to prevent or retard the development of the arthropathy is infusion of factor concentrate. This can be carried out ‘on demand’ in the event of a bleed, or ‘prophylactically’ by regular continuous treatment [1,2,4–6]. Regardless of the treatment regimen, the outcome must be monitored using medical, financial and quality of life protocols. Orthopaedic and radiological scoring systems are used to evaluate joint function and arthropathy [7–10].

The most widely used methods for radiological assessment of haemophilic arthropathy are the Arnold–Hilgartner scale [10] and the Pettersson score [8]. These methods are based on conventional X-ray technology (radiography). They were outlined several decades ago, and at that time were useful for evaluating the degree of joint destruction in patients with haemophilia. To date, improvement of therapy has led to the need for more refined systems to monitor the less advanced changes in arthropathy. With radiography, it is difficult to evaluate soft tissues, therefore the pathological process is often underestimated. By comparison, magnetic resonance imaging (MRI) has tomographic capability and provides good soft tissue contrast, thus it has the potential to provide more detailed information [11–25]. In recent years, several investigators have presented MRI scoring systems for evaluation of haemophilic arthropathy [19,21–25]. The first of these, and so far the most extensively developed, is the Denver scale designed by Nuss and colleagues [21,25]. Both this scale and the Arnold–Hilgartner scale employ a ‘progressive’ strategy, in which the most severe change determines the score. Notwithstanding, there is still no consensus about the best way to use MRI to score haemophilic arthropathy.
In this work, we applied both the Denver scale and a new scheme (the European scoring method) to evaluate MRI of the ankles of children with haemophilia. Our objective was to present an MRI score that offers increased sensitivity to minor progression of the arthropathy and also allows different components of the pathology to be monitored separately. In addition, we wanted to investigate the intra- and interobserver agreement of MRI readings.

Patients and methods

In a collaborative project within the European Paediatric Network for Haemophilia Management (PedNet; see Acknowledgements), 39 ankles in 28

Fig. 1. Left ankle of 12-year-old boy with moderate haemophilia A. By consensus, this joint was assigned a Denver score of 10 and a European score of 8(2:3:3). (a) Sagittal standard spin echo (SE) T1-weighted sequence (TR 418 ms, TE 12 ms, slice thickness 3 mm, field of view 100 mm, matrix 184 \( \times \) 256, two acquisitions). The image shows irregularity/erosion of subchondral cortex of the tibia and talus (small arrows), as well as hypertrophic synovia with haemosiderin depositions (haemosiderin is black, large arrows). (b) Sagittal turbo spin echo (TSE) T2-weighted sequence (TR 4200 ms, TE 96 ms, echo train 7, slice thickness 3 mm, field of view 100 mm, matrix 154 \( \times \) 256, two acquisitions). The scan detects effusion (fluid is bright, arrow). (c) Sagittal fat suppressed 3D gradient echo (GE) sequence (TR 50 ms, TE 11 ms, flip angle 40°, slice thickness 1.5 mm, field of view 100 mm, matrix 228 \( \times \) 256, one acquisition). The image shows destruction of talar cartilage (joint fluid in defect is bright, arrow). (d) Coronal short tau inversion recovery (STIR) sequence (TI 150 ms TR 4500 ms, TE 60 ms, slice thickness 4 mm, field of view 113 mm, matrix 198 \( \times \) 256, three acquisitions). The scan reveals subchondral cyst in tibia (arrow).
children 4–16 years (mean 10 years) of age with severe and moderate haemophilia were investigated by MRI at nine paediatric haemophilia centres in Europe.

The participating investigators were given a four-sequence protocol designed for a 1.5 T VISION MRI system (Siemens, Erlangen, Germany) (Fig. 1), and they were asked to adjust their MRI equipment to achieve images that were as similar as possible.

Working independently, two experienced musculoskeletal radiologists (BL and HP) evaluated the images twice, and each time classified the investigated ankles according to both the previously published Denver score [21,25] (Table 1) and the new scoring method designed in the present work (the European score; Table 2). Final classification was agreed upon by consensus.

Statistical analysis

Intra- and interobserver agreement were estimated for the two scoring methods. The A, e, s and h components of the European score were analysed separately. Both unweighted and weighted kappa values were calculated; a kappa value is basically a measure of agreement between observations. Values indicate the degree of agreement as follows: 1.00–0.81 very good; 0.80–0.61 good; 0.60–0.41 moderate; 0.40–0.21 fair; and <0.20 poor. A weighted value differs from an unweighted value in that it is influenced to a lesser extent by small variations in paired observations. Used in combination, weighted and unweighted values indicate not only the strength of agreement, but also the proportions of small and large differences.

Results

The new European MRI scoring scheme for evaluation of haemophilic arthropathy was designed in the format A(e:s:h) (Table 2). The A component of the score (maximum value 16) is calculated as the sum of values for subchondral cysts (maximum 6), irregularity/erosion of subchondral cortex (maximum 4) and chondral destruction (maximum 6). In these three categories of arthropathic changes, a number of statements (six, four and six, respectively), are evaluated as to whether they are true or false, and each true statement adds one point to the A component. The factors e, s and h represent effusion/haemarthrosis, hypertrophic synovia and haemosiderin deposition, and they are evaluated separately according to a five-grade scale (0–4) as follows: 0 is normal; 1 is equivocal; 2, 3 and 4, respectively, designate that the degree of pathology is small, moderate and large.

Table 1. The Denver magnetic resonance imaging score in which different stages of the pathology are classified in relation to the most severe finding (assigned the maximum score of 10).  

<table>
<thead>
<tr>
<th>Effusion/haemarthrosis</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Normal joint</td>
<td>Small</td>
<td>Moderate</td>
<td>Large</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Synovial hyperplasia/haemosiderin</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Small</td>
<td>Moderate</td>
<td>Large</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cyst/erosion</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 cyst or partial surface erosion</td>
<td>&gt;1 cyst or full surface erosion</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cartilage loss</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt;50%</td>
<td>≥50%</td>
</tr>
</tbody>
</table>

Table 2. The European magnetic resonance imaging score given in the format A(e:s:h).*

<table>
<thead>
<tr>
<th>Subchondral cysts (part of A)</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present in at least one bone</td>
<td>Absent</td>
<td>Equivocal</td>
<td>Small</td>
<td>Moderate</td>
<td>Large</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Irregularity/erosion of subchondral cortex (part of A)</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present in at least one bone</td>
<td>Absent</td>
<td>Equivocal</td>
<td>Small</td>
<td>Moderate</td>
<td>Large</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chondral destruction (part of A)</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full-thickness defect in at least one bone</td>
<td>Absent</td>
<td>Equivocal</td>
<td>Small</td>
<td>Moderate</td>
<td>Large</td>
</tr>
</tbody>
</table>

*Maximum score 16(4:4:4); explained in detail in the ‘Results’.

Table 3 shows the individual European scores of the 39 investigated ankle joints, and the corresponding Denver scores, all as were arrived at by consensus.

© 2004 Blackwell Publishing Ltd
The results shown in Table 4 indicate that intraobserver agreement was good or moderate for both scoring methods (unweighted kappa values were 0.63/0.70 for the Denver score and 0.77/0.64, 0.58/0.77, 0.71/0.59 and 0.64/0.63, respectively, for the A, e, s and h components of the European score). Interobserver agreement was somewhat poorer (unweighted kappa values 0.56/0.38 for the Denver score and 0.51/0.42, 0.54/0.56, 0.71/0.35 and 0.34/0.29 for the components of the European score). For both intra- and interobserver agreement, the weighted kappa values were higher than the unweighted values.

### Discussion

Management of haemophilia has improved dramatically over the past few decades, mainly due to increased supply and higher quality of factor concentrate [2–7,9,19]. However, arthropathy is still a major concern for people with haemophilia, and better methods are needed to evaluate subtle arthropathic changes in these patients in order to minimize joint damage and optimize the cost-effectiveness of treatment.

Radiological assessment of disease processes can be accomplished using various imaging techniques and different strategies to evaluate the images that are obtained. In haemophilic arthropathy, radiography is limited in that it cannot satisfactorily visualize important features of the condition, for instance effusion/haemarthrosis, synovial hypertrophy, cartilage destruction and subchondral bone changes. By comparison, such changes are more clearly discerned by MRI [11–25], thus this technique has the potential to provide additional information.

According to the Arnold–Hilgartner radiographic scale [10], a disease passes through specific stages that have a characteristic appearance in the recorded images. Staging is categorized in relation to the most...
The methods employing this strategy have been called ‘progressive’ [21], and they are especially useful if the various radiographic stages correlate with different clinical stages that are important for making therapeutic decisions. A disadvantage of this approach is that progression of the disease is documented only if a more advanced stage is observed; in other words, progressive changes occurring in the current or earlier stages will not influence the results. The Pettersson radiographic score uses another strategy based on summation of specific imaging findings [8]. In such methods, which have been referred to as ‘additive’ [21], all findings influence the assessment, and the sensitivity for detecting progression of the disease is greater. Superior sensitivity is particularly important when evaluating haemophilic arthropathy, which is why the World Federation of Hemophilia recommends use of the Pettersson score.

The Denver scoring scheme using MRI (Table 1) and the Arnold–Hilgartner scale using radiography both employ the same progressive type of strategy. The MRI-technique resolves more constituents of the arthropathy than radiography does, and consequently it allows the disease to be divided into a larger number of stages. In comparison with the Arnold–Hilgartner scale, the Denver score provide better resolution and enables detection of early changes.

In comparison with the Denver score, the European MRI scoring system described in this paper (Table 2) divides each category of pathological change into a greater number of steps. Furthermore, this scheme uses an additive strategy to increase the resolution, and it separates different pathological components. To date, the destructive changes in bone and cartilage (A) that are associated with haemophilic arthropathy are considered to be irreversible, whereas other components of the disease, for instance effusion/haemarthrosis (e) and synovial hypertrophy (s), may regress. The European scoring system monitors such changes separately. Hemosiderin deposition (h) is also considered separately, although this may not represent a reversible change, and measurement of this feature at present has no known practical significances. A haemarthrosis is an acute episode promoting the development of arthropathic changes, but is not a change of interest to include in a score aiming at a quantitative estimate of the arthropathy itself. Usually, it is apparent from the clinical situation if a haemarthrosis is present, and scoring should be avoided during such episodes. However, with MRI differentiation between effusion and haemarthrosis is often difficult, why these entities are combined in one component (e).

In MRI examinations, it is essential that the images obtained are of sufficiently good quality for diagnostic purposes. It is usually necessary to increase the imaging time to improve the quality of the images, which can limit the number of joints that can be evaluated on a single occasion. We used a four-sequence protocol, and the imaging of one joint took about 30 min. This strategy was feasible, because only one or two joints were investigated in each patient. In other diseases, as childhood arthritis, intravenous gadolinium contrast media has been used to facilitate the discrimination of synovial hypertrophy and effusion. However, such advocacy renders the procedure invasive and more costly. We avoided this similar to other authors designing MRI scoring methods for haemophilic arthropathy [19,21–25].

The majority of the patients we studied had low MRI scores (Table 3). More precisely, 26 of the 39 investigated joints had a Denver score of 4 or lower, which reflects our objective of assembling patients exhibiting early arthropathic changes in order to develop a suitable method of evaluating such changes. Considering the results of the Denver method, only one of the 13 remaining joints had a score in the intermediate range, and the other 12 joints were given scores of 9 or 10. These results indicate the effect of the progressive strategy of this method: as cartilage destruction evolves the Denver score shifts to the highest levels. By comparison, the additive strategy of the European scoring scheme allows further discrimination of MRI findings.

Considering the results of our statistical evaluation (Table 4), it should be noted that the A, e, s and h components of the European score were analysed separately, and the analysis was based on a limited number of investigations. Nevertheless, the results obtained using the two different scoring methods are essentially similar. Interobserver agreement was good or moderate, which indicates that it is possible to achieve consistent evaluation. The values representing interobserver agreement were lower than those for intraobserver agreement, because film reading is to some extent subjective and different readers use slightly different criteria to make diagnoses. Such differences should be reduced as much as possible. Measures for this can be specified definitions of subjective terms such as small, moderate, etc., and distribution of case examples for comparison. A first book with this aim is recently published [25]. Our results show such measures are important and that the need for special training of readers even applies to experienced musculoskeletal radiologists. We also found that weighted kappa values were higher than the unweighted results, which reflects the existence of
few large differences between evaluations. This observation is interesting, because it implies that, besides being able to provide mean values for groups of patients in a research context, MRI scoring may also allow clinical monitoring of individual patients.

The role of the European- and other MRI scores in management of haemophilia will be addressed in future works. Our results in this work are encouraging and suggest that MRI can be a valuable tool that will help to improve management of patients with haemophilia and optimize the outcome of treatment. A number of questions remain to be answered. Can radiologists be trained to perform standardized reading? How do MRI scores correlate with clinical data? Which MRI scoring method and imaging protocol should be used?

In conclusion, MRI can be useful for semiquantitative assessment of haemophilic arthropathy, if the procedure is carefully performed using an adequate protocol and the images are evaluated by specially trained radiologists.

Acknowledgements

This work was supported by grants from the Swedish Research Council (no. 13493), the University of Lund (ALF) and funds administered by Malmö University Hospital. This project required an international multicentre effort, and it could not have been conducted without the contributions of members of the European Paediatric Network for Hemophilia Management (PedNet) and their cooperating radiologists. The PedNet acknowledge the support given to the group from Bayer AG. Special thanks to the following participants: Drs M. Siimes and P. Tervahartiala, Helsinki, Finland; Drs J. Tusell and A. Castellette, Barcelona, Spain; Drs S. Clayssens and C. Baunin, Toulouse, France; Drs H. Chambost and P. Petit, Marseille, France; Drs E. Scheibel and K.-E. Jensen, Copenhagen, Denmark; Drs S. Aronis and A. Papadopoulos, Athens, Greece; Drs F. Hill and K. Johnson, Birmingham, UK; Drs P. Petrini and T. Finnbogason, Stockholm, Sweden; Dr C. Widebaeck, Malmö, Sweden. We are also grateful to Drs R. Nuss and R. Kilcoyne of the University of Colorado (Denver, CO, USA) for previous achievements in the field of magnetic resonance imaging of haemophilic arthropathy and for valuable personal communications.

References


18 Nuss R, Kilcoyne RF, Geraghty S, Wiedel J, Manco Johnson M. Utility of magnetic resonance imaging for...


22 Funk MB, Schmidt H, Becker S et al. Modified magnetic resonance imaging score compared with orthopaedic and radiological scores for the evaluation of haemophilic arthropathy. *Haemophilia* 2002; 8: 98–103.

