Association between US and MRI scores and correlations with urinary C-terminal telopeptide Type II collagen levels for early detection of hemophilic arthropathy of the knee

M Prasetyo1, D Gatot2*, A B M Tulaar3, J Pandelaki1, S Bardosono4, L Sukrisman5, A M T Lubis6, S G Malik7 and M Ilyas8

1Department of Radiology, Faculty of Medicine, Universitas Indonesia, Jakarta, 10430, Indonesia
2Department of Pediatrics, Faculty of Medicine, Universitas Indonesia, Jakarta, 10430, Indonesia
3Department of Physical Medicine and Rehabilitation, Universitas Indonesia, Jakarta, 10430, Indonesia
4Department of Nutrition, Faculty of Medicine, Universitas Indonesia, Jakarta, 10430, Indonesia
5Department of Internal Medicine, Faculty of Medicine, Universitas Indonesia, Jakarta, 10430, Indonesia
6Department of Orthopedic and Traumatology, Faculty of Medicine, Universitas Indonesia, Jakarta, 10430, Indonesia
7Eijkman Molecular Biology Institute, Faculty of Medicine, Universitas Indonesia, Jakarta, 10430, Indonesia
8Department of Radiology, Faculty of Medicine, Universitas Hasanuddin, Makassar, 90245, Indonesia

*E-mail: djagat_ps@yahoo.com

Abstract. Assessment of knee hemophilic arthropathy (HA) requires objective measures. There is no consensus on a preferable ultrasonography (US) scoring system, whereas a reliable scoring system for magnetic resonance imaging (MRI) was developed by the International Prophylaxis Study Group of the World Federation of Hemophilia (IPSG-WFH). In this cross-sectional study, a new US scoring system for early detection of knee HA was developed based on 25 peer-reviewed articles published between 1999 and 2015; furthermore, its associations with the MRI scoring system and urinary C-terminal telopeptide type II collagen (CTX-II) levels were investigated. In total, 27 children with severe hemophilia A were recruited. Early HA was confirmed using radiography as Arnold–Hilgartner stages 0–II. Knee MRI and US were scored using the MRI IPSG-WFH and new US scoring systems; urinary CTX-II levels were measured using enzyme-linked immunosorbent assay. Correlations were identified using Spearman test. The US scoring system included joint effusion, synovial hypertrophy and hypervascularization on power Doppler US, hemosiderin deposition, and cartilage damage. The US and MRI scores showed a moderate correlation; the US score and urinary CTX-II levels showed no correlation. The new US score can be used as an alternative to MRI for assessing early-stage knee HA.
1. Introduction

Hemophilic arthropathy (HA) is a progressive joint disease that occurs in approximately 70–80% of hemophiliacs due to recurrent bleeding, especially in the large joints, such as the knee joint [1]. Data from the National Hemophilia Integrated Service Center of Cipto Mangunkusumo General Hospital recorded 180 cases of HA, which included 115 cases of severe hemophilia A [2]. Progression of HA can be prevented by the administration of clotting factor therapy as early as possible after the onset of hemarthrosis. As HA can be a lifelong illness, the cost of treatment becomes very expensive, as clotting factor therapy alone accounts for 50–90% of the total health care cost for hemophilia, depending on the type and severity of the condition [3] which is a considerable problem especially in developing countries. For hemophiliacs, advanced HA drastically decreases quality of life [4-6].

Preventing the progression toward severe HA is a primary goal of hemophilia treatment. Therefore, an evaluation method to objectively assess anatomical changes of the joints and the severity of joint HA is necessary to ensure proper management [7]. Radiological examinations by radiography, ultrasonography (US), and magnetic resonance imaging (MRI) can provide such information. Radiography is used more often for the assessment of advanced stage HA, while US and MRI are more reliable for the assessment of the early stage of HA. Although MRI is the best method, equipment is not widely available because of the high expense and the need to sedate children. Thus, US is more commonly used because it is easier to implement for the evaluation of early-stage HA [8-14]. Nevertheless, standardization of measurable assessment methods is still needed.

The Arnold–Hilgartner staging system is used to describe articular changes with disease progression, and the Pettersson scoring system is used to grade the severity of HA [15,16]. An MRI scoring system has been developed by the International Prophylaxis Study Group of the World Federation of Hemophilia (IPSG-WFH) [7,17,18]. Although other scoring systems have been developed, there is currently no consensus, and accurate scoring is technically difficult and thus requires much experience. Moreover, as compared with MRI, US is insufficient to assess all parts of the cartilage due to the limited penetration of sonographic waves that cannot reach the cartilage in the mid-joint region [11].

Therefore, a new US scoring system is needed that is easier to apply and does not require extensive experience, but is yet in line with the standard MRI scoring system. To overcome the limitations of US for the assessment of cartilage, the proposed scoring system was also correlated with urinary C-terminal telopeptide type II collagen (CTX-II) as a biochemical marker that is sensitive to cartilage damage [19,20]. The development of this US scoring system is expected to be very useful for routine evaluations of hemarthrosis in hemophiliacs. This scoring system can be used to assess the efficacy of clotting factor therapy to determine an appropriate therapeutic strategy as well as for future research and clinical trials [10,11,13].

2. Methods

This cross-sectional study was conducted using medical records from February 2015 to October 2016 archived by the Department of Radiology and the Eijkman Molecular Biology Institute, Faculty of Medicine, Universitas Indonesia. A new US scoring formulation was conducted based on 25 related peer-reviewed articles published from 1999 to 2015.

After confirming the validity of the proposed scoring system, patient records that included radiographical data confirming the early stage of HA of the knee (Arnold–Hilgartner classification of 0–II) were included for analysis. On the day of the US examination, morning urine was collected and stored at −20 °C. Musculoskeletal US examinations of the knee joint were performed with an Esaote® US system (Esaote SpA, Genoa, Italy) equipped with an 11–13 MHz linear frequency transducer. All US examinations were performed by a researcher and a radiologist with more than 10 years of experience in the field of US and MRI. Disparities in interpretations of the results were resolved by agreement. MRI examinations of the knee were performed after US using a Magnetom Avanto® 1.5 Tesla MRI system (Siemens Healthineers, Erlangen, Germany) equipped with a special knee coil to acquire T1-weighted image spin echo, T2-weighted image turbo spin echo image, short tau inversion recovery, and fat-
saturated three-dimensional gradient recalled echo sequences. Assessment of US and MRI scores was performed after the collection of all data. The US score was determined with the proposed scoring system, while the MRI score was determined with the IPSG-WFH scoring system. The urinary concentration of CTX-II was determined with a commercial enzyme-linked immunosorbent assay (ELISA) kit (Urine CartiLaps®; Immunodiagnostic Systems, The Boldons, UK). The samples were diluted to the detection range of the ELISA kit. The optical densities of the samples were determined with a spectrophotometer and corrected to urinary creatinine levels. The Spearman Rho correlation test was performed to identify correlations between the US and MRI scores, as well as between the US score and the corrected urinary CTX-II level.

3. Results

The proposed US scoring system based on peer-reviewed studies retrieved from the literature was developed to assess anatomic changes in the early stage of HA of the knee: i.e., synovial fluid accumulation (joint effusion/hemarthrosis), synovial hypertrophy, synovial hypervascularization, hemosiderin deposition, and cartilage defects. With this scoring system, the degree of the anatomic change was simply assigned a value of 0–3. To reduce complexity, the components of synovial hypertrophy, hemosiderin deposition, and cartilage defect were classified as mild or severe. The degree category was combined with the degree of severity because the primary goal was early detection, so the distinction was whether the changes were mild or not. Synovial fluid accumulation was classified as mild, moderate, or severe, as more detailed information was needed to assess the therapeutic response. Synovial hypervascularization, as determined by power Doppler US (PDUS), was only indicated by the presence of a Doppler signal. Thus, the maximum score of the five components combined was 10 points (Table 1).

Joint effusion was pathologically confirmed when fluid accumulation was found to be thicker than 4 mm. Since the expansion of joint fluid was always found in the supra and para patella recesses, the division of degrees was focused to the compartment, i.e., mild = fluid accumulation limited to the supra or para patella recesses; moderate = fluid accumulation in the supra patella recess and either the medial or lateral para patella recess; and severe = fluid accumulation in all supra, medial, and lateral para patella recesses.

| Component                                      | Definition                                                                 | Category | Score | Information                                |
|------------------------------------------------|---------------------------------------------------------------------------|----------|-------|--------------------------------------------|
| Synovial fluid accumulation (joint effusion)  | Anechoic structure with a thickness of ≥4 mm, with or without debris filling the supra patella and/or para patella | Small    | 1     | Detected in the supra patella or para patella |
|                                                |                                                                           | Medium   | 2     | Appeared in the supra patella and medial or lateral para patella |
|                                                |                                                                           | Large    | 3     | Appeared in the supra patella, the medial and lateral para patella |
| Synovial hypertrophy                           | Clearly bounded hypoechoic curvilinear structure that was a genu joint capsule, especially in the supra and para patella recesses | Small    | 1     | Thin, not easily visualized*                |
|                                                |                                                                           | Medium/large | 2  | Thick, easily visualized**                 |

Table 1. Proposed US Scoring System of HA of the Knee.
Table 1. Continue

| Component                        | Definition                                                                 | Category           | Score | Information                                |
|----------------------------------|---------------------------------------------------------------------------|--------------------|-------|--------------------------------------------|
| Synovial hypervascularization (PDUS) | Color-coded signal of blood flow to the thickened synovium wall          | Present            | 1     | Positive on PDUS                          |
| Hemosiderin deposition           | Not clearly bounded hypoechoic structure like a cloud, could be detected in the synovium cavity or attached to the synovium wall | Small              | 1     | Vague, not easily visualized*              |
|                                  |                                                                           | Medium/large       | 2     | Thick and easily visualized**              |
| Cartilage thickness defect       | Trochlea femoris cartilage displaying irregularity/defect; estimated at less than or more than 50% of the visible surface | Small              | 1     | Present, <50% surface                     |
|                                  |                                                                           | Medium/large       | 2     | Present, >50% surface                      |

Total scores 0–10

Information:
*Not easily visualized: visible after performing certain transducer techniques, such as compression, angulation, and dynamic study techniques.
**Easily visualized: the abnormality was visible without a particular transducer technique.

The synovium was more easily visualized in the supra patella recess, which under normal circumstance is not visualized. Synovial hypertrophy was considered mild when the synovium appeared slightly thickened in the supra or para patella recesses, which requires a special US transducer technique (dynamic study) for confirmation. Synovial hypertrophy was considered severe when the synovium was very thick and visible, without the need for a dynamic study. Quantitative measurements were avoided because a variation in handling of the transducer of only a few millimeters would greatly affect the measurement results. In addition, the location of the measurement could also be influential, as the thickening of the synovium may not be homogeneous across the synovium walls (Figure 1).

Synovial hypervascularization was confirmed by thickening of the synovium wall on PDUS, which indicated the onset of an immune response to hemarthrosis, in both acute and chronic conditions of exacerbation. Hemosiderin deposition, in the form of unclear binding of hypoechoic structures to the thickened synovium wall, is a consequence of blood degradation. Deposition was considered to be mild when found in a minimal amount, which requires a dynamic study for clear observation, and severe when deposition was easily visualized without a dynamic study or with synovial hypertrophy.

Cartilage defects, as observed by US, were limited to the trochlea femoris only. Considering that the purpose of the examination was early detection, the severity of the defect was determined with regard to the extent of damage to the cartilage surface due to either erosion or defect. Damage to <50% of the visible surface was considered mild, while damage to >50% of the visible surface was considered severe.

Data were obtained from the medical records of 80 boys with severe hemophilia by the Integrated Hemophilia Service Team of the RSCM. After adjusting for the study criteria, 40 subjects were selected for correlation analysis between the US score and urinary CTX-II level. Meanwhile, 27 subjects were recruited consecutively according to the study criteria for correlation analysis between the US and MRI scores (Figure 2).
The cohort for correlation analysis between the US and MRI scores included 27 subjects with an average age of 10.5 (range, 6–16) years, median MRI IPSG-WFH score of 1 (range, 0–6) point, and median USG score of 1 (range, 0–8) point. The results of the Shapiro–Wilk normality test showed that the data were abnormally distributed ($p < 0.05$). The Spearman test showed a significant positive correlation between the new US score and the MRI IPSG-WFH score ($p = 0.012$, one-tailed) with moderate correlation strength ($R = 0.431$). The adjusted $R^2$ value obtained with the multiple regression test was 0.868.

Figure 1. Knee MRI of Score 6 on HA Subject.

Figure 2. Early-Stage HA with a US Score of Eight. Effusion Fluid (F) Appeared in the Supra Patella and the Medial-Lateral Para Patella Recesses (A–C). Synovial Hypertrophy was Easily Visualized by Observation of Hypervascularization and Hemosiderin Deposition by PDUS (A). No Cartilage Erosion was Observed (D).
The study cohort for correlation analysis between the new US score and urinary CTX-II level included 40 subjects with an average age of 10 (range, 6–17) years, a median knee US score of 1 (range, 0–8) point, and a median corrected urinary CTX-II level of 31,851 (range, 6801–889,772) ng/mmol. The distribution of urinary CTX-II levels was abnormal and there was a very large range of urinary CTX-II values, with extreme values found in three subjects. The results of the Spearman correlation test showed that there was no significant correlation between the new US score and corrected urinary CTX-II level ($p = 0.360$, one-tailed), with very low correlation strength ($R = 0.058$). The absence of a correlation with urinary CTX-II level still required further analysis, as the CTX-II level had greatly varied, suggesting the contribution of other factors, such as the severity of arthropathy in other joints, treatment compliance, and/or the presence of new recurrent bleeding before the examination.

4. Discussion

Here, we report the development of a new scoring system based on the early stage of HA pathophysiology, as well as an evaluation of pre-published scoring systems, especially in aspects of complexity, degree of difficulty, and relevance. It is important to note that a complicated scoring system for the assessment of HA will have an impact on the low availability of services, as it seems to be the case with MRI. Classification with previous scoring systems was quite complex with four levels of abnormality for a minimal anatomical change. In addition, osteochondral components are overelaborated, which according to the researchers’ view, are susceptible to overdiagnosis or overstaging due to the anisotropy factor introduced by the US transducer, whereas the true osteochondral changes can be best visualized only on radiographs [13,21]. This occurs because the previous scoring system was developed to assess the early to advanced stages of joint changes. Accordingly, a new scoring system was developed with a focus on early anatomical joint changes and the use of categorization criteria that was easier to learn. However, this scoring system was also expected to remain in line with the other standardized scoring systems, such as the IPSG-WFH version of the MRI scoring system. There are five components assessed in this US scoring system, i.e., joint effusion, synovial hypertrophy, synovial hypervascularization (PDUS), hemosiderin deposition, and cartilage defect. The minimum total score is 0 and the maximum is 10.

All subjects in this study had severe hemophilia with recurrent joint bleeding in the knee joints, as well as the elbow and ankle joints. The scores of the 27 subjects who were included for correlation analysis were abnormally distributed, which was expected considering that the focus of this study was early detection with US, so the MRI and US scores of most samples were low. In addition, the ages of patients were abnormally distributed, which could be explained by the lack of sedation and requirement of a certain level of cooperation during the MRI examinations. Thus, the cohort was limited to children aged >10 years.

The results of Spearman correlation analysis found a positive correlation between the total US score and total MRI score ($p = 0.012$, $R = 0.431$). Some previous studies on US scoring systems failed to investigate correlations with the MRI score. The scoring system according to Melchiorre et al. [10] obtained a positive correlation to the Pettersson score radiographically, with a correlation coefficient of 0.374–0.440. The correlation strengths found in this study might have been influenced by the differences in the number of assessed components, as well as the different categories of change. Nevertheless, as the proposed scoring system is easier to be understood and to apply, this medium degree of correlation strength shows that the US scoring system is useful and relevant for early detection of HA cases in Indonesia (Figure 3).
The proposed US score shows not only the magnitude of the score, but also the components that affect the total score. The following score components were included in addition to the total score: the effusion component (E) with a score of 0–3, synovial hypertrophy (S) with a score of 0–2, synovial hypovascularization on PDUS (P) with a score of 0–1, hemosiderin deposition (H) with a score of 0–2, and cartilage defect (K) with a score of 0–2. For example, a maximum total score of 10 will have the configuration score $E_3S_2P_1H_2K_2$. In this situation, the fluctuating components (E, P, and H) can be differentiated from the progressive components (S and K). If the S and K components reach the maximum scores, it can be expected that HA is not in the early stage, and more sensitive examinations with either radiography or MRI are needed for assessment of advanced disease. Thus, if it is not possible to perform MRI for one reason or another, the combination of US and radiography should prove to be very helpful, as changes to the soft tissue can be visualized by US and changes to the bone by radiography.

Multiple regression analysis showed a good correlation strength ($R = 0.943$, adjusted $R^2 = 0.868$), which meant that the new US score contributed 86% to the MRI IPSG-WFH score, suggesting that the new US score can be used as an alternative to MRI. However, in actual application, serial US scores will be more useful. If the US score persists in serial US examinations, it can be assumed that there has been a change in the osteochondral component and MRI or a radiographic examination is needed to assess

\[ E_3S_2P_1H_2K_2 \]

**Figure 3.** Implementation Algorithm of the New Scoring System.
changes to the osteochondral component. As long as the US score continues to fluctuate and is less than the maximum value, it can be assumed that there is no advanced stage of HA and the evaluation with US was sufficient.

There was no significant correlation between the US score and urinary CTX-II level ($p = 0.360$) due to the number of joints involved, since the US score assessed one knee joint only while the obtained CTX-II level was representative of cartilage degradation throughout the body. It was initially thought that as the largest joint, cartilage degradation of the knee joint would most likely influence the high concentration urinary CTX-II. In fact, almost all subjects had a history of swelling of the elbow, knee, and ankle joints with varying levels of urinary CTX-II. Considering the radiographic picture that shows a low degree of damage to both knees, unknown factors may have contributed to the damage to the elbow and ankle joints.

This suggestion was reinforced by additional analysis of differences in CTX-II levels by age, the number of joints involved, the number of target joints, and the type of target joints. Based on the Mann–Whitney test, no significant difference in urinary CTX-II level was found between patients aged ≤10 and >10 years ($p = 0.713$), >6 and ≤6 joints involved ($p = 0.255$), >6 and ≤6 target joints ($p = 0.079$), as well as the type of target joints (knee joint vs. not the knee joint) ($p = 0.628$). Thus, it can be assumed that other factors influence the high levels of CTX-II, such as the severity of damage to the six major joints, a previous history of bleeding, and/or daily activities, although these data were not obtained in this study. Similarly, the knee joint is larger than the elbow and ankle joints, so it was thought to be the most influential to the high levels of CTX-II, although there were actually no significant differences.

As a limitation of this study, the validity and reliability of the proposed scoring system have not yet been determined; so further studies are warranted. In addition, some data, such as treatment compliance, history of joint bleeding in the last 2 weeks, and the severity of HA in large joints other than the knee, were not included, which may have affected the correlation between the US score and urinary CTX-II level.

5. Conclusions
A new US scoring system has been developed for evaluation of early-stage HA of the knee. There was a positive correlation with moderate correlation strength between the US score and the IPSG-WFH version of MRI score for the early detection of HA, while there was no correlation between US score and urinary CTX-II level.

References
[1] World Federation of Hemophilia 2012 Guidelines for the management of hemophilia 2nd ed. (Montreal: World Federation of Hemophilia).
[2] Himpunan Masyarakat Hemofilia Indonesia 2015 Data registrasi penyandang hemofilia di Indonesia. Tidak dipublikasi.
[3] Rocha P, Carvalho M, Lopes M and Araújo F 2015 Costs and utilization of treatment in patients with hemophilia. *BMC Health Serv. Res.* 15 484.
[4] Lee C A, Berntorp E and Hoots K 2010 *Textbook of hemophilia* (Chichester, West Sussex, UK; Hoboken, NJ: Wiley-Blackwell).
[5] Gater A, Thomson T A and Strandberg-Larsen M 2011 Haemophilia B: impact on patients and economic burden of disease. *Thromb. Haemost.* 106 398.
[6] Astermark J, Dolan G, Hilberg T, Jimenez-Yuste V, Laffan M, Lassila R, Lobet S, Martinoli C and Perno C F 2014 Managing haemophilia for life: 4th Haemophilia Global Summit. *Haemophilia* 20 1.
[7] Lundin B, et al 2012 An MRI scale for assessment of haemophilic arthropathy from the International Prophylaxis Study Group. *Haemophilia* 18 962.
[8] Zukotynski K, Jarrin J, Babyn P S, Carcao M, Pazmino-Canizares J, Stain A M and Doria AS et
al 2007 Sonography for assessment of haemophilic arthropathy in children: a systematic protocol. *Haemophilia* 13 293.

[9] Klukowska A, Czyzny Z, Laguna P, Brzowski M, Serafin-Krol M and Rokicka-Milewska R 2001 Correlation between clinical, radiological and ultrasonographical image of knee joints in children with haemophilia. *Haemophilia* 7 286.

[10] Melchiorre D, Linari S, Innocenti M, Biscoglio I, Toigo M, Cerinic M M and Morfini M 2011 Ultrasound detects joint damage and bleeding in haemophilic arthropathy: a proposal of a score: sonographic score in haemophilic arthropathy. *Haemophilia* 17 112.

[11] Martinoli C, et al 2013 Development and definition of a simplified scanning procedure and scoring method for Haemophilia Early Arthropathy Detection with Ultrasound (HEAD-US). *Thromb. Haemost.* 109 1170.

[12] Sierra Aisa C, et al 2014 Comparison of ultrasound and magnetic resonance imaging for diagnosis and follow-up of joint lesions in patients with haemophilia. *Haemophilia* 20 e51.

[13] Doria A S 2010 State-of-the-art imaging techniques for the evaluation of haemophilic arthropathy: present and future. *Haemophilia* 16 107.

[14] Doria A S, et al 2015 Diagnostic accuracy of ultrasound for assessment of hemophilic arthropathy: MRI correlation. *Am. J. Roentgenol.* 204 W336.

[15] Arnold W and Hilgartner M 1977 Hemophilic arthropathy: current concepts of pathogenesis and management. *J. Bone Ji. Surg. Am.* 59 287.

[16] Pettersson H, Ahlberg A and Nilsson I 1980 A radiologic classification of hemophilic arthropathy. *Clin. Orthop.* 149 153.

[17] Doria A S, Lundin B, Kilcoyne R F, Babyn P S, Miller S, Nuss R, Rivard G, Stephens D and Pettersson H 2005 Reliability of progressive and additive MRI scoring system for evaluation of haemophilic arthropathy in children: Expert MRI Working Group of the International Prophylaxis Study Group. *Haemophilia* 11 245.

[18] Doria A S, Babyn P S, Lundin B, Kilcoyne R F, Miller S, Rivard G E, Moineddin R, Pettersson H 2006 Reliability and construct validity of the compatible MRI scoring system for evaluation of haemophilic knees and ankles of haemophilic children. Expert MRI working group of the international prophylaxis study group. *Haemophilia* 12 503.

[19] Jansen N W D, Roosendaal G, Lundin B, Heijnen L, Mauser-Bunschoten E, Bijlsma J W, Theobald M and Lafeber F P 2009 The combination of the biomarkers urinary C-terminal telopeptide of type II collagen, serum cartilage oligomeric matrix protein, and serum chondroitin sulfate 846 reflects cartilage damage in hemophilic arthropathy. *Arthritis Rheum.* 60 290.

[20] Van Vulpen L F D, van Meegeren M E R, Roosendaal G, Jansen N W D, van Laar J M, Schutgens R E G, Mastbergen S C and Lafeber F P 2015 Biochemical markers of joint tissue damage increase shortly after a joint bleed; an explorative human and canine in vivo study. *Osteoarthritis Cartilage.* 23 63.

[21] Ng W H, Chu W C W, Shing M K, Lam W W M, Chik K W, Li C K, Li C K and Ling S C 2005 Role of imaging in management of hemophilic patients. *Am. J. Roentgenol.* 184 1619.