

1 Intra- and inter-observer reliability of arthroscopic quantification of chondropathy and synovial
2 changes in horses with osteochondral fragments in the tibiotarsal joint.
3 Authors: Nielsen, J.V., DVM
4 Institution: Ansager Veterinary Hospital
5 Address reprint to: Jon V. Nielsen, MRCVS, Ansager Veterinary Hospital, Gartnerhaven 5, 6823
6 Ansager, Denmark. Mail: jonveddingnielsen@ansagerdyrehospital.dk
7 Keywords: Arthroscopy, scoring, inter-observer, intra-observer, tibiotarsal joint.
8 (Word count: 3.778)
9

1

2 **Summery**

3 *Reasons for performing study: Arthroscopic evaluation of joints is a well established diagnostic*
4 *procedure but establishment and validation of different intra-articular scoring systems have not*
5 *been performed in equine practice.*

6 *Objectives: To investigate the intra- and inter-observer reliability of different scoring systems and*
7 *to evaluate the effect of a training session on the inter-observer reliability.*

8 *Methods: Ten horses were selected from a population of horses that underwent arthroscopic*
9 *removal of OCD in the tibiotarsal joint with simultaneous video digital recording. The ten horses*
10 *were selected for findings that represent different stages of severity of chondropathy in the*
11 *tibiotarsal joint. The arthroscopic recordings were evaluated in a pre-training assessment, a*
12 *training session and post-training assessment.*

13 *Results: There were no significant differences in all the scoring systems between the surgeons at the*
14 *pre-training assessment. All scoring systems except the SFA score showed significant difference*
15 *between the surgeons in the post-training assessment. This study showed acceptable reliability of*
16 *the SFA score, VAS average and VAS (chondropathy, osteophytes and synovium).*

17 *Conclusions and potential relevance: This study indicates a possible use of the SFA score, VAS*
18 *average and VAS (chondropathy, osteophytes and synovium).*

19

20 **Introduction**

21 Through the last two decades arthroscopy has reached status as the best diagnostic tool in the
22 evaluation of joints. The possibility of a detailed visual examination of the arthroscopic visible
23 cartilage provides surgeons with a diagnostic tool in the establishment of the most accurate
24 prognosis possible in cases of joint related lameness's. The technique allows simultaneously intra-

1 articular surgery with debridement and removal of intra-articular fragments or osteochondral
2 lesions, with a concurrently prognostic evaluation. Though there have been an extensive
3 advancement in other diagnostic modalities such as digital x-rays, computer tomography,
4 ultrasonography and magnetic resonance imaging, arthroscopy still remains the most sensitive
5 imaging technique in detecting the cartilage surface defects characteristic of osteoarthritis (Adams
6 *et al.* 1991; McIlwraith *et al.* 2005; Spiers *et al.* 1993). Several scoring systems are being used in
7 the recording of human cartilage damage and synovial changes. The grading system of equine
8 cartilage damage, which is currently being used, is a categorical classification with four groups:
9 grade I, softening and swelling of the cartilage without loss of surface continuity; grade II,
10 superficial fissuring of the cartilage surface; grade III, deep fissuring of the cartilage surface, with
11 the fissures reaching the subchondral bone, and erosion/ulceration of the cartilage without exposure
12 of bone; grade IV, erosions down to bone. These four categories are extrapolated from human
13 medicine where the grading system was proposed by Beguin and Locker as a variant of the
14 classification proposed by Outerbridge (Beguin and Locker 1983; Outerbridge 1961). This grading
15 system only considers the depth of the lesion but neither the size nor location. Noyes and Stabler
16 introduced a classification system to human knee arthroscopists which evaluated the four important
17 variables: 1) the articular surface, 2) the extent or depth of the involvement, 3) the diameter of the
18 lesion and 4) the location of the lesion (Noyes and Stabler 1989). Klashman *et al.* proposed a
19 variant of this classification in which grade I is divided into two subcategories grade Ia; intact but
20 dull articular surface, grade Ib swelling and/or softening, grade II into IIa and IIb superficial or deep
21 fissures or erosions, respectively, and grade IIIa and IIIb exposed bone without or with bone
22 cavitation, respectively (Klashman *et al.* 1995). The French Society of Arthroscopy have proposed
23 two scoring systems: the SFA Scoring System (Soci t  Fran aise d'Arthroscopie) using a visual
24 analogue scale (0-100 mm) and the SFA Grading System, which is a categorical evaluation with

1 four pathological categories (Ayraral *et al.* 1994; Dougados *et al.* 1994). The American College of
2 Rheumatology (ACR) proposed a scoring system in which the chondropathy is recorded on a
3 continuous scale from 0-350 (Ayraral *et al.* 1998; Klashman *et al.* 1995). In human medicine, studies
4 have validated the use of different scoring systems (Table 1) (Ayraral *et al.* 1998; Dougados *et al.*
5 1994; Klashman *et al.* 1995). A review of the equine literature shows only one study using and
6 validating a scoring system – the SFA grading system. Brommer *et al.* validated the SFA
7 arthroscopic grading system in the evaluation of the correlation between arthroscopic findings and
8 gross pathological findings (Brommer *et al.* 2004). This study showed no significant difference
9 between two surgeons, but unfortunately the SFA arthroscopic grading system is not applicable in
10 the equine joints when the primary lesion is osteochondrosis. The ultimate arthroscopic scoring
11 system includes severity, size and location of cartilage defects, but also includes an evaluation of
12 the non-cartilage structures. Furthermore the scoring system has to be easy to use with high
13 accuracy and a high intra- and inter-observer correlation. The aim of the present study was to
14 investigate the applicability of different human arthroscopic scoring systems on horses, and hereby
15 evaluate the intra- and inter-observer reliability. The scoring systems were applied to horses
16 undergoing arthroscopic removal of osteochondral fragments in the tibiotarsal joint. The hypothesis
17 was tested that there is good intra-observer and inter-observer correlation when using the different
18 scoring systems.

19

20 **Materials and methods**

21 *Sample population*

22 Ten horses were selected from a population of 200 horses that underwent arthroscopic removal of
23 OCD with simultaneous video digital recording in the period from Mai to November 2007 at

1 Ansager Veterinary Hospital, Denmark. The ten horses were selected for OCD in the tibiotarsal
2 joint and findings representing different stages of severity of chondropathy.

3 *Observers*

4 The study included three arthroscopists with 20, 12 and 5 years of experience in arthroscopy. The
5 recordings were obtained by the different surgeons and assessments were performed afterwards by
6 all three surgeons.

7 *Arthroscopy and recording*

8 Under general anesthesia with the horse placed in dorsal recumbency the affected joints were
9 inspected and palpated for the presence of joint effusion and fibrous thickening of the joint capsule.
10 Arthroscopy of the tibiotarsal joints was performed using the dorsomedial approach as described by
11 McIlwraith (McIlwraith *et al.* 2005). The joints were effused with isotonic saline and the
12 arthroscope (4 mm, 30 degree)¹ was inserted through a skin incision in the dorsomedial
13 outpouching just dorsal to the center of the dorsomedial outpouching. The arthroscopic findings
14 were recorded on a Lemke LCAP Motion digital image capture system² with the recordings
15 performed from lateral to medial.

16 *Scoring systems*

17 A form was created to include the contents of the different scoring systems:

18 1) Overall Assessment of Chondropathy in the tibiotarsal joint is evaluated using a 100 mm visual
19 analogue scale (VAS) with 0 indicating no chondropathy and 100 indicating the most severe
20 chondropathy (Ayrar *et al.* 1993). This evaluation is done for every joint surface; distal dorsolateral
21 tibia, lateral trochlea, distal intermediate tibial ridge, medial trochlea and medial maleoli. An
22 average VAS of the joint surfaces is subsequently calculated.

1 2) Overall Assessment of Degeneration of the tibiotarsal joint of using a 100 mm VAS. This
2 assessment includes in addition to the Overall Assessment of Chondropathy also the formation of
3 osteophytes.

4 3) The SFA Scoring System (Dougados *et al.* 1994). The chondropathy is recorded as percentage of
5 the entire cartilage surface with grading of the depth according to the classification of chondropathy
6 proposed by Beguin and Locker (Beguin and Locker 1983). The SFA score is obtained as follows:

7 $A = \text{size (\%)} \text{ of grade I lesions} \times 0.14$

8 $B = \text{size (\%)} \text{ of grade II lesions} \times 0.34$

9 $C = \text{size (\%)} \text{ of grade III lesions} \times 0.65$

10 $D = \text{size (\%)} \text{ of grade IV lesions} \times 1.00$

11 $\text{SFA score} = A+B+C+D$

12 The SFA score is a continuous variable ranging from 0 to 100, where 0 indicates no chondropathy
13 and 100 indicates the most severe chondropathy.

14 4) The ACR Scoring System (Dougados *et al.* 1994) records depth and size where depth is recorded
15 by the use of the classification proposed by Klashman *et al.* (1995). The sizes of the different lesion
16 grades (grade I to IV) are evaluated on a 100 mm VAS in which 0 indicates absence of
17 chondropathy and 100 the most severe chondropathy. The ACR score is obtained as follows:

18 $A = \text{size (mm)} \text{ of grade Ia lesions} \times 1$

19 $B = \text{size (mm)} \text{ of grade Ib lesions} \times 1.5$

20 $C = \text{size (mm)} \text{ of grade IIa lesions} \times 2.0$

21 $D = \text{size (mm)} \text{ of grade IIb lesions} \times 2.5$

22 $E = \text{size (mm)} \text{ of grade IIIa lesions} \times 3.0$

23 $F = \text{size (mm)} \text{ of grade IIIb lesions} \times 3.5$

24 $\text{ACR score} = A+B+C+D+E+F$

- 1 The ACR score is a continuous variable ranging from 0 to 350, where 0 indicates no chondropathy
2 and 350 indicates the most severe chondropathy with total exposure of bone.
- 3 5) Overall Assessment of Synovial structures of the tibiotarsal joint of using a 100 mm VAS. This
4 assessment does not include chondropathy.
- 5 6) Overall Assessment of Osteoarthritis in the tibiotarsal joint of using a 100 mm VAS. This
6 assessment includes chondropathy, osteophytes and synovial structures.
- 7 7) Size of the osteochondral fragment was assessed using description in mm and a categorical
8 variable ranging from 1-3 with 1 being the largest fragment.
- 9 8) Osteochondral lesions in the cartilage bigger than the fragment was assessed using description in
10 mm.
- 11 9) An Assessment of Osteoarthritis of the tibiotarsal joint with a categorical variable ranging from
12 1-3 with 1 being the highest degree of changes.
- 13 10) Finally the prognosis for recovery to complete athletic use was evaluated as excellent, fair,
14 reserved or poor.

15

16 *Study design*

17 The recordings of the ten horses were analyzed 1 week (pre-training evaluation) prior to and 8
18 weeks (post-training evaluation) after a training and evaluation session.

19 Pre-training evaluation: The recordings were delivered to the three surgeons for evaluation. This
20 was done after a brief explanation of the scoring forms without information on interpretation or
21 description of the lesions to be visualized. The recordings were independently reviewed and
22 analyzed by the surgeons based on prior experience with arthroscopy.

23 Training evaluation: The recordings were reviewed during a session with all three surgeons
24 together. Evaluation of every recording was performed and discrepancies between the surgeons

1 were discussed accordingly. Agreement was reached in every case resulting in a session evaluation
2 of every case.

3 In between evaluations: Recordings were performed as part of a study performed on cartilage-
4 derived retinoic acid-sensitive protein (CD-RAP) also referred to as melanoma-inhibitory-activity
5 protein (MIA) in horses with different joint conditions. The recordings were performed and
6 afterwards analyzed independently by the three surgeons.

7 Post-training evaluation: The three surgeons reviewed and analyzed the recordings of the same ten
8 horses 8 weeks later. The recordings were as in the pre-training evaluation independently reviewed
9 and analyzed for the third time.

10

11 *Statistical analysis*

12 Bartlett's test for inequality and homogeneous of variances was performed. When the assumptions
13 of normality and homogeneity were confirmed the null hypothesis of equal means were tested in a
14 nonparametric ANOVA repeated measurement test³; H_0 no difference between groups ($P < 0.05$
15 considered significant). This was done for each of the different continuous response variables
16 between the surgeons and the pre- and post-training evaluations. Univariate linear regression
17 analysis was performed to determine the correlation between the surgeons' evaluations and between
18 the pre- and post-training evaluations⁴. This was done with a predetermined intercept of zero.

19 Pearson's correlation coefficients (r), F-values and parameter estimates with P-values and
20 confidence intervals were calculated.

21

22 **Results**

23 *Descriptive statistics*

1 This study included 10 horses all undergoing arthroscopic surgery. Two horses were excluded from
2 the study due to incomplete pre-training evaluation forms. The remaining 8 horses came from a
3 mixed population with 5 Danish Warmbloods, 2 Frisians and 1 standardbred trotter. The mean age
4 was 2.8 years. There was a history available on 6 of the 8 horses with no history of lameness by any
5 of the horses. Two horses had x-rays taken due to routine pre-purchase examination and 4 horses
6 had x-rays taken due to joint effusion. Six horses had clinical symptoms with all of these having
7 joint effusion and only one having a fibrous thickened joint capsule.

8
9 Significant differences between horses were found in all outcome variables listed in Table 2 which
10 were expected since the eight horses for this part of the study were selected by an overall
11 preliminary assessment, to represent varying degrees of cartilage damage and synovial
12 inflammation. According to all scoring systems horse #1 had the most severe changes and horse #2
13 and #3 moderate changes and #4, #5, #6, #7 and #8 with low degree of cartilage damage and
14 synovial inflammation. These findings were consistent with the athletic prognosis specified to these
15 horses (Table 2). Pre-training assessment showed a relatively large variation in all scoring systems
16 with a greater variation between the ACR scores (Table 2). According to Table 2 there were no
17 obvious signs of differences between the variances of the scoring systems between the different
18 degrees of the joint pathology. The post-training assessment showed a higher variation in almost all
19 scoring systems due to surgeon 1 consistently having higher values in all scoring systems except the
20 ACR score (Table 3). There was no significant difference in the scoring systems between the
21 surgeons in the pre-training assessment, but significant differences in VAS, SFA and ACR scores
22 were found in the post-training assessment with surgeon 1 being an outlier (Table 3).

23

1 There was no significant difference between the scores of the different scoring systems and the
2 prognosis applied. In the same way there was no significant difference between the sizes of the
3 fragments and the prognosis applied.

4
5 The correlation between the surgeons in the pre-training assessment and the consensus agreed in the
6 training session showed extensive to exemplary consistency (approximately 0.80) in most scoring
7 systems (Table 4a). The correlation coefficients of all scoring systems were all significantly
8 different from 0 except for the ACR score recorded by surgeon 2. The correlation coefficient of the
9 VAS synovium by surgeon 1 showed moderate consistency (Table 4a). Surgeon 2 and 3 had
10 significantly lower SFA scores in the pre-training assessment compared to the training assessment.
11 Surgeon 1 and 2 had significantly lower ACR values in the pre-training assessment compared to the
12 training session (Table 4a).

13
14 The correlation between the post-training and the training session decreased to a value not
15 significantly different from 0 in the case of surgeon 1 in all scoring systems except the SFA and
16 ACR score (Table 4b). This was also seen at the VAS synovium score performed by surgeon 2
17 (Table 4b). In the post-training assessment all scoring systems had one or more surgeons who had
18 significantly higher or lower scores compared with the training assessment (Table 4b). Apart from
19 surgeon 1 and the VAS synovium score, there was no decrease in correlation between pre- and post-
20 training correlation with the training assessment. The ACR score showed a marked increase in
21 correlation with the training in the post-training assessment (Table 4b).

22
23 The intra-observer correlation was in this study “moderate” to “exemplary” for all scoring systems
24 with the highest correlation by the SFA score and lowest by the VAS synovium (Table 3c). The

1 lowest intra-observer correlation was seen by surgeon 1 in all scoring systems except for the VAS
2 synovium which has a generally low intra-observer correlation. Significant higher VAS synovium
3 scores were found by all surgeons post-training. Besides a lower correlation surgeon 1 had
4 significant higher scores in all scoring systems.

5
6 The only significant differences found between the surgeons in the pre-training assessment were the
7 difference in the characterisation of the synovial membrane and hereby the presence of the different
8 characteristics of the synovial membrane (Table 3). Significant differences were found between the
9 surgeons when evaluating the presence of petechiation in the synovial membrane, and whether the
10 changes found were localised or diffuse distributed in the joint. This finding was consistent in the
11 post-training assessment with the addition of significant differences in the presence of hyperaemia
12 and deposits of haemosiderin (Table 3).

14 **Discussion**

15 The different scoring systems evaluated in this study are extrapolated from human medicine where
16 the systems have been developed and validated particularly for knee arthroscopy (Ayrar *et al.*
17 1998). Although the anatomic structures involved in arthritis and osteoarthritis are the same, the
18 selection criteria for arthroscopy in human and equine patients are different. The most frequent
19 indication in equine medicine is removal of osteochondral fragments which was the indication by
20 all horses in this study. In human medicine the most frequent indication is joint trauma or
21 rheumatoid osteoarthritis and this difference will cause a difference in the degree of changes and a
22 difference in the presence of age-related changes between equine and human joint pathology.
23 Establishment and validation of scoring systems in equine medicine is required in order to carry out
24 multicenter studies and/or in the interest of a uniform medical terminology.

1

2 All the scoring systems in this study were able to demonstrate the difference between the horses
3 which were selected to represent different degrees of chondropathy. This means that all the scoring
4 systems are able to differentiate between different degrees of chondropathy, which is the first
5 criteria of a useful scoring system. It was not possible to show any significant difference between
6 the scores of the different scoring system and the prognosis applied. This was probably due to the
7 limited number of horses in this study, but also due to a lack of divergence between the prognoses
8 of the selected horses.

9

10 The universal goal of all scorings or evaluations is to achieve complete agreement between
11 surgeons. The consensus achieved on the recordings in the training session was therefore considered
12 to be the golden standard for the pre- and post-training evaluation. This was only possible because
13 complete agreement in every assessment were reached. This study design have been used in human
14 studies not only to evaluate intra- and inter-observer correlation within institutions but also in
15 evaluations of scoring systems discussed and studied at international meetings of rheumatologists
16 (Ayrál *et al.* 1998).

17

18 The pre-training evaluations were performed with a prior brief explanation of the use of the scoring
19 systems (VAS, SFA and ACR) but without any discussion of the interpretation or description of
20 pathology to be visualized. Under these circumstances the scoring systems showed “exemplary”
21 (≥ 0.80) and “extensive” (0.70-0.79) correlation with the training assessment, with the exception of
22 the ACR score showing “moderate” (0.60-0.69) and “minimal” (< 0.60) correlation. The very low
23 correlation of surgeon 2 in the ACR score was explained by a specific misunderstanding of this
24 scoring system. This study design was created with the purpose to study inter-observer reliability of

1 the scoring systems when the surgeons applied these for the first time, but also to eliminate bias
2 from different experience with the scoring systems. This study was carried out with three
3 arthroscopists with 5-20 years of experience in arthroscopy. Though this experience is enough to
4 get familiar with the normal variation and most of the pathologic changes, some of the inter-
5 observer variation could be explained by the surgeons' discrepancies of findings and their
6 significance. Arthroscopic evaluation of joints is subjective and will always be biased by the
7 surgeons opinion and experience, and cause some inter-observer variance when compared. The use
8 of categorical variables will reduce the subjective aspect and thereby minimize inter-observer
9 variance, but this will create a suboptimal scoring system as it will not evaluate all three description
10 parameters; depth, size and location. The size estimations were in this study based solely on
11 subjective measurement which can be improved by the use of direct measurement with intra-
12 articular size measuring equipment. This would make the scoring systems more quantitative and
13 therefore more objective. The quantitative evaluation in general (VAS, SFA score and ACR score)
14 has the advantage in comparison to the categorical to better evaluate deterioration of cartilage over
15 time, which will be desirable in different study designs evaluating disease progression and treatment
16 regimes. Longitudinal studies are warranted to evaluate the scoring systems ability to detect
17 cartilage changes over time.

18
19 It was chosen to place the post-training evaluation 8 weeks after the training session. This was done
20 in order to avoid bias caused by the surgeons' ability to remember the pre-training and training
21 evaluations. In the relapsed period the surgeons evaluated ten recordings using the same setup of
22 scoring systems. This was done as part of another study (Berg, L. C. 2007) and to create the most
23 realistic circumstances of how scoring systems will be used in a daily basis at equine hospitals
24 performing arthroscopy. The intra-observer correlation was in this study "moderate" to "exemplary"

1 for all scoring systems, with the highest correlation by the SFA and lowest by the VAS synovium.
2 This indicates that all the scoring systems can be used by surgeons repeatedly in the reassessing of
3 cases and comparison of cases performed by the same surgeon. The significant higher values in
4 post-training compared to pre-training assessment cause some concern but this can perhaps be
5 explained by increased experience or the consensus achieved at the training session. The study
6 design itself can cause intra-observer variation due to the interactive discussion of the cases and
7 compromises performed by surgeons to reach the consensus. If this was the case, it would have
8 been expected that the inter-observer differences were not significant. The inter-observer
9 differences in the post-training assessment were significant but explained by surgeon 1 in all cases,
10 except the ACR score which was explained by surgeon 2. This difference could not be related to the
11 surgeon's experience in arthroscopy.

12
13 The lowest intra-observer correlation was seen by surgeon 1 in all scoring systems, except for the
14 VAS synovium which had a general low intra-observer correlation. Significant higher VAS
15 synovium scores were found by all surgeons post-training. Surgeon 1 had besides a lower
16 correlation also significant higher scores in all scoring systems. Apart from surgeon 1 and the VAS
17 synovium score there was no decrease in correlation between pre- and post-training correlation with
18 the training assessment. The ACR score showed a marked increase in correlation with the training
19 in the post-training assessment which indicates the importance of a training session when this
20 scoring system is applied. The significant difference in ACR scores between surgeons in the post-
21 training assessment does not advocate for this system as a reliable scoring system.

22
23 The evaluation of the synovium showed the lowest inter- and intra-observer correlation combined
24 with significant differences between surgeons in the pre- and post-training assessment. This

1 indicates a lack of reliability when surgeons assess the synovial structures. This study showed
2 acceptable reliability of the SFA score, VAS average and VAS (chondropathy, osteophytes and
3 synovium).

4
5 Ayral *et al.* (1998) performed a human study equivalent to this study with five patients classified
6 according to the ACR classification as knee osteoarthritis (Ayral *et al.* 1998). Ayral *et al.* showed
7 improved and acceptable inter-observer reliability, with a scoring system using 100 mm VAS
8 producing best inter-observer reliability. Differences between human studies (Ayral *et al.* 1998;
9 Dougados *et al.* 1994) and the findings in this study can be explained by the differences in the
10 selection criteria of patients undergoing arthroscopy.

11
12 This pilot study indicates a possible use of the SFA score, VAS average and VAS (chondropathy,
13 osteophytes and synovium), but to choose the most accurate and reliable scoring system requires
14 further studies. This will include a larger study of the inter-observer reliability and the
15 establishment of a key-score (recordings with a consensus) to adjust the surgeons scores. Further
16 studies include usability in the evaluation of other joints i.e. metacarpophalangeal joint and stifle
17 joints, longitudinal studies of the scoring systems ability to detect changes, studies regarding
18 correlation with clinical symptoms and studies investigating the correlation between scoring
19 systems and the macroscopic changes found by dissection.

20
21 Intra-observer reliability is a paramount criterion of a scoring system, since reproducibility is the
22 most vital aspect of all evaluations with or without scoring systems. The inter-observer reliability of
23 scoring systems is necessary in order to carry out multicenter studies and to achieve uniform
24 medical terminology. Further and larger studies are necessary to evaluate the inter-observer and

- 1 intra-observer reliability of these scoring systems with particular interest in the SFA score, VAS
- 2 average and VAS (chondropathy, osteophytes and synovium).

¹ Karl Storz Veterinary Endoscopy, Goleta, USA. www.ksvea.com.

² Lemke Vision, W.O.M. World of Medicine AG, Berlin, Germany. www.lemkevision.com.

³ Epi-Info version 3.3.2, February 9, 2005. Atlanta, Georgia USA. www.cdc.gov/epiinfo.

⁴ Microsoft® Office Excel 2003.

3

4

5

6

7

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37

References

- Adams, M.E., Li, D.K., McConkey, J.P., Davidson, R.G., Day, B., Duncan, C.P., and Tron, V. (1991) Evaluation of cartilage lesions by magnetic resonance imaging at 0.15 T: comparison with anatomy and concordance with arthroscopy. *J.Rheumatol.* 18, 1573-1580.
- Ayral, X., Dougados, M., Lustrat, V., Bonvarlet, J.P., Simonnet, J., Poiraudou, S., and Amor, B. (1993) Chondroscopy: a new method for scoring chondropathy. *Semin.Arthritis Rheum.* 22, 289-297.
- Ayral, X., Gueguen, A., Ike, R.W., Bonvarlet, J.P., Frizziero, L., Kalunian, K., Moreland, L.W., Myers, S., O'Rourke, K.S., Roos, H., Altman, R., and Dougados, M. (1998) Inter-observer reliability of the arthroscopic quantification of chondropathy of the knee. *Osteoarthritis.Cartilage.* 6, 160-166.
- Ayral, X., Gueguen, A., Lustrat, V., Bahuaud, J., Beaufils, P., Beguin, J., Bonvarlet, J.P., Boyer, T., Coudane, H., Delaunay, C., and . (1994) [Proposal of a simplified arthroscopic score of cartilage lesions of the knee (simplified SFA score)]. *Rev.Rhum.Ed Fr.* 61, 97-99.
- Beguin, J. and Locker, B. (1983) Chondropathie rotulienne. *2^{ème} Journée d'Arthroscopie du Genou, Lyon, 16-17 septembre* 1, 89-90.
- Berg, L. C., (2007) Equine MIA / CD-RAP in horses with osteochondrosis. Unpublished work. Department of Basic Animal and Veterinary Sciences, Faculty of Life Sciences, Grønnegårdsvej 7, 1870 Frederiksberg C, Denmark.
- Brommer, H., Rijkenhuizen, A.B., Brama, P.A., Barneveld, A., and Weeren, P.R. (2004) Accuracy of diagnostic arthroscopy for the assessment of cartilage damage in the equine metacarpophalangeal joint. *Equine-Veterinary-Journal* 36, 331-335.
- Dougados, M., Ayral, X., Lustrat, V., Gueguen, A., Bahuaud, J., Beaufils, P., Beguin, J.A., Bonvarlet, J.P., Boyer, T., Coudane, H., and . (1994) The SFA system for assessing articular cartilage lesions at arthroscopy of the knee. *Arthroscopy* 10, 69-77.
- Klashman, D., Ike, R., Moreland, L., Skovron, M.L., and Kalunian, K. (1995) Validation of an OA data report form for knee arthroscopy. *Arthritis and rheumatism / Official journal of the American Rheumatism Association* 38(Suppl),
- McIlwraith, C.W., Nixon, A.J., Wright, I.M., and Boening, K.J. 2005 Diagnostic and surgical arthroscopy of the tarsocrural (tibiotarsal) joint, Third edition, Mosby Elsevier, London. pp 269-306.
- Noyes, F.R. and Stabler, C.L. (1989) A system for grading articular cartilage lesions at arthroscopy. *Am J.Sports Med.* 17, 505-513.
- Outerbridge, R.E. (1961) The etiology of chondromalacia patellae. *J.Bone Joint Surg.Br.* 43, 752-757.

1 Spiers, A.S., Meagher, T., Ostlere, S.J., Wilson, D.J., and Dodd, C.A. (1993) Can MRI of the
2 knee affect arthroscopic practice? A prospective study of 58 patients. *J.Bone Joint Surg.Br.*
3 **75, 49-52.**
4
5