

# Relationship of Radiographic Osteoarthritis Severity with Treatment Outcomes after Imaging-Guided Knee Injections: A Prospective Outcomes Study

## Zusammenhang zwischen Kniearthrose-Schweregrad und Behandlungserfolg nach bildgesteuerter Infiltration: Eine prospektive Kohortenstudie

### Authors

Ivan Miletic<sup>1</sup>, Christoph Agten<sup>2</sup>, Reto Sutter<sup>2</sup>, Christian Pfirrmann<sup>2</sup>, Cynthia Peterson<sup>2</sup>

### Affiliation

- 1 Chiropractic, Dr. Jörg Häusler private practice, Luzern, Switzerland
- 2 Radiology, Balgrist University Hospital, Zurich, Switzerland

### Key words

knee osteoarthritis, radiography, treatment outcomes, injections, intraarticular

received 10.01.2017

accepted 22.05.2017

### Bibliography

DOI <https://doi.org/10.1055/s-0043-112337>

Published online: 26.6.2017 | Fortschr Röntgenstr 2018; 190: 134–143 © Georg Thieme Verlag KG, Stuttgart · New York, ISSN 1438-9029

### Correspondence

Ivan Miletic  
Chiropraktik, Praxis Dr. Häusler, Inseliquai 8, 6005 Luzern, Switzerland  
Tel.: ++41/76/4894067  
[ivan.miletic@chirosuisse.ch](mailto:ivan.miletic@chirosuisse.ch)

### ZUSAMMENFASSUNG

**Ziel** Um vorhersagen zu können, welche Kniearthrose-Patienten ein positives Behandlungsergebnis nach einer intraartikulären Corticosteroidinjektion erwarten können, sind die Ziele dieser Studie 1: Den Schweregrad der Arthrosen mittels unterschiedlichen Einteilungssystemen zu bestimmen und diese mit den Ergebnissen von Infiltrationen zu vergleichen. Die dabei verwendeten Einteilungen sind das Kellgren und Lawrence (KL) System mit 5 und mit 3 Graden für Kniearthrose, das Osteoarthritis Research Society International (OARSI) Einteilungssystem für Gelenkspaltverschmälerung sowie Messungen des effektiven Gelenkspaltes; 2: Die Reliabilität dieser Einteilungssysteme miteinander zu vergleichen.

**Material und Methoden** Knie-Röntgenbilder von 117 Patienten, welche eine Corticosteroidinjektion erhalten haben, wurden von zwei Radiologen unabhängig voneinander und ohne

die Ergebnisse der Infiltrationen zu kennen evaluiert. Die Evaluationen beinhalteten die Einteilung des Schweregrades der Arthrose mit den KL5-, KL3-, OARSI-Systemen sowie der Messungen des effektiven Gelenkspaltes. Die klinischen Vergleichsparameter wurden erfasst mittels numerischer Schmerzskala vor der Behandlung und zusammen mit dem „Patient's global impression of change score“ nach einem Tag, einer Woche und einem Monat nach der Intervention. Die Proportionen derjenigen Patienten innerhalb der verschiedenen Einteilungen, welche über eine Besserung berichteten, wurden verglichen mittels Chi-square-Test. Die logistische Regression wurde genutzt um festzustellen, welche Befunde ein positives Ergebnis voraussagen können. Die Übereinstimmung der beiden Radiologen wurde verglichen mittels Kappa-Koeffizient und dem Intraclass correlation coefficient (ICC).

**Ergebnisse** Patienten mit OARSI Grad 2 hatten zu allen Zeitpunkten die höchsten Proportionen, welche über eine Verbesserung berichteten, signifikant am Tag 1 ( $p = 0,004$ ). Kein Zusammenhang zwischen Behandlungsergebnis und Arthrose Grad konnte festgestellt werden mit den Einteilungen nach KL5, KL3 oder effektiven Gelenkspaltmessungen. Patienten mit OARSI Grad 2 berichteten 8 Mal häufiger über eine Verbesserung am Tag 1 ( $p = 0,024$ ). Effektive Gelenkspaltmessungen hatten die beste Reliabilität ( $ICC = 0,812 - 0,882$ ), gefolgt vom OARSI-Einteilungssystem.

**Schlussfolgerungen** Patienten mit Kniearthrose Grad OARSI 2 (34 – 66 % Gelenkspaltverschmälerung) haben bessere Ergebnisse nach Corticosteroidinjektionen nach einem Tag, einer Woche und einem Monat. Die Reliabilität der OARSI-Einteilung war besser als diejenige von KL5 und KL3.

### Kernaussagen

- Arthrose-Einteilungen nach OARSI hatten eine bessere Reliabilität als die Einteilungen nach KL5 oder KL3.
- OARSI Grad 2 hatte die besten Behandlungsergebnisse.
- Weder die Einteilungen nach Kellgren und Lawrence noch die effektiven Gelenkspaltmessungen waren verbunden mit einer klinischen Verbesserung nach erfolgter Infiltration.

## ABSTRACT

**Introduction** To be able to predict which patients are more likely to have a positive treatment outcome, the purpose of this study is 1: To compare outcomes after intra-articular corticosteroid knee injections with the Kellgren and Lawrence (KL) 5 and 3 grading systems for knee osteoarthritis, the Osteoarthritis Research Society International (OARSI) grading system and actual joint space measurements; and 2: To compare the reliability of these grading systems.

**Materials and Methods** Knee radiographs of 117 patients who received intra-articular corticosteroid injections were independently evaluated by two radiologists blinded to the outcome. Evaluation included the KL5, KL3, OARSI systems and actual joint space widths. The numerical rating scale for pain was collected at baseline and along with the Patient's Global Impression of Change on day 1, in week 1 and in month 1. The number of 'improved' patients was compared between the OA grades using the Chi-square test. Logistic regression determined which findings were predictive for improvement. Agreement was assessed using Kappa statistics and the intra-class correlation coefficient (ICC).

**Results** Patients with OARSI grade 2 reported the highest rates of 'improvement' at all time points, which was significant on day 1 ( $p = 0.004$ ). No relationship with improvement was found with KL5, KL3 or actual joint space measurements. Patients with OARSI grade 2 were 8 times more likely to report improvement on day 1 ( $p = 0.024$ ). Reliability was best for joint space measurements (ICC = 0.812–0.882), followed by the OARSI.

**Conclusion** The OARSI for joint space narrowing grade 2 (34–66% narrowing) was linked with a better outcome on day 1 with trends in week 1 and month 1. The reliability of the OARSI was better than the KL5 or KL3 systems.

### Key Points

- OARSI grading of OA had better reliability than KL3 or KL5.
- OARSI grade 2 was related to a better treatment outcome.
- Neither KL grades nor joint space measurements were related to improvement.

### Citation Format

- Miletic I, Agten C, Sutter R et al. Relationship of Radiographic Osteoarthritis Severity with Treatment Outcomes after Imaging-Guided Knee Injections: A Prospective Outcomes Study. *Fortschr Röntgenstr* 2018; 190: 134–143

## Introduction

Intra-articular corticosteroid injections are a popular therapeutic intervention for painful joints and are being widely used to treat the various types of rheumatological and osteoarthritic joint disorders. In fact, there seems to be a good short-term benefit with respect to pain reduction, which lasts up to three weeks [1–5], and the long-term use of repetitive injections has been shown to be safe and effective in relieving symptoms for some patients [6]. A recent systematic review by Jüni et al., which looked at the efficacy of intra-articular corticosteroid injections in knee OA in 2015, states that important benefits are unclear due to the low quality of evidence [7] and that the use of steroids in knee OA therefore remains controversial [4, 7]. For clinicians, it would be very important to know which patients respond well to this treatment and which do not.

Since plain radiographs are inexpensive, widely available, do not require any special facilities and are often the first and only tool to evaluate degenerative changes of joints, it would be of interest to determine if abnormal findings on these radiographs could be used as a predictive factor for treatment outcome. Since there are different radiological grading systems with incongruences, it is also important to know the strengths and weaknesses of these different grading systems. Two of the most widely used grading systems for OA today are the Kellgren and Lawrence grading system and the more recently developed Osteoarthritis Research Society International atlas criteria (OARSI) [8, 9]. The KL system was published in 1957, and adopted by the World Health Organization in Rome in 1961 as the accepted gold standard for cross-sectional and longitudinal epidemiological studies [10].

While the KL system defines OA severity in five grades (0 = normal to 4 = severe) using a combination of osteophyte and joint space narrowing severity, the later developed OARSI atlas uses semi-quantitative separate scoring for osteophytes and joint space narrowing (grading 0–3) [8].

Studies in which plain radiographs were compared with arthroscopy support the need for more sensitive grading, because the plain radiographs significantly underestimated the extent of degenerative changes, especially regarding cartilage abnormalities in osteoarthritic joints [11, 12]. Even patients with no radiographic findings of OA had significant articular cartilage degeneration within the femorotibial joint in arthroscopy [13]. Femorotibial joint space narrowing reflects cartilage loss in knee OA [12, 14–16] and is more sensitive and therefore more accurately assesses progression of OA than osteophyte formation [8, 17–20]. Thus, it would make sense to have a closer look at joint space measurements and the OARSI for joint space narrowing to see if these could be better used as predictive factors for outcome after steroid injections. This is particularly relevant as the OARSI for joint space narrowing alone would be a very easy tool to use in daily practice.

Therefore, the objectives of this study are: 1) To compare outcome after intra-articular corticosteroid injections into the knee with the KL 5 and 3 grading systems for knee osteoarthritis, the OARSI grading system for joint space narrowing and actual joint space measurements; 2) to compare the reliability of these different grading systems by assessing the inter-rater reliability.

## Materials and methods

### Patients

This is a retrospective evaluation of knee radiographs from patients in a prospective outcomes cohort study. Data of 117 consecutive patients who received an imaging-guided therapeutic (corticosteroid plus anesthetic) intra-articular knee injection in the period of 4/14/2009 to 2/10/2014 at the radiology department of this specialized orthopedic university hospital, with weight-bearing anterior-posterior (AP) and recumbent lateral radiographs taken within 6 months of the injections and who returned an outcomes questionnaire by mail were included in this study. Hospital and cantonal ethics approval (EK-12/2009) was obtained prior to the start of this study and all patients gave their informed consent. After the injection, the patients were given an outcomes questionnaire and were asked to complete it at one day, one week, and one month after the injection. The follow-up outcomes questionnaire was given to the patients by the radiological technologist with a stamped and addressed envelope with instructions to return the completed questionnaire one month after the injection.

### Knee injection procedure

The injections were all performed by radiologists from a specialized orthopedic university hospital. Under sterile conditions (3 × disinfection, sterile gloves, mask, cover) and with fluoroscopic control, the involved knee was punctured with a 22-gauge needle. Arthrography was performed with 2 ml Iopamiro 300 (Iopamidol). Infiltration of 1 ml Triamcort (Triamcinolone 40 mg/ml) and 5 ml Rapidocaine 2% (Lidocaine 20 mg/ml) was then performed. Intra-articular distribution of the injected contrast material and, therefore, correct needle placement were documented with a radiograph (► Fig. 1).

### Patient data collection and outcomes

Before the injection, each patient's pain level was recorded using the numerical rating scale (NRS), where 0 means no pain and 10 is the worst imaginable pain. This served as the baseline NRS score. Fifteen minutes after the injection, the NRS was measured again. In the questionnaires, which the patients were asked to send back after one month, the NRS for one day, one week and one month after the injection as well as the Patient's Global Impression of Change (PGIC) scale were obtained. The PGIC consists of a scale from 1 to 7, where 1 means much better, 2 better, 3 slightly better, 4 no change, 5 slightly worse, 6 worse and 7 much worse [21–23].

In this study the PGIC scale was dichotomized such that only scores of 1 and 2 were included as 'improvement', with all other responses considered as 'not improved'. Similarly, scores of 5–7 were considered 'worse' and all other scores were 'not worse'. This dichotomization is identical to other studies using the PGIC scale [21–23]. 'Improvement' was the primary outcome measure. The NRS change score was calculated by subtracting the one month NRS score from the baseline NRS score.



► Fig. 1 Knee radiograph showing correct needle position and contrast distribution.

► Abb. 1 Knie-Röntgenbild mit Kontrastmittelverteilung indiziert korrekte Nadelpositionierung.

### Radiographic Evaluation

The severity of OA on the weight-bearing AP knee radiographs of the patients who received a therapeutic knee injection were read and classified independently and blinded to the clinical outcomes by a skeletal radiology fellow and by a radiologist. The OA classification was done using three different grading systems for knee OA. The Kellgren and Lawrence system with 5 grades, a simplified version with only 3 grades, as well as the osteoarthritis research society international grading system (OARSI) for medial and lateral femorotibial joint space narrowing (► Table 1) were used [16, 17, 24–26]. Joints were scored based on the compartment with the worst radiographic findings. Examples of different grades are shown in ► Fig. 2a–c. The medial and lateral joint spaces were measured electronically on the hospital PACS system as shown in ► Fig. 3.

### Statistical analysis

The primary outcome measure was the proportion of patients reporting 'improvement' compared to the various OA classifications. The Chi-square test was used for this analysis. The proportion of patients reporting 'worsening' was also compared for the various OA classification systems using the Chi-square test (secondary outcome).

Logistic regression was done to see which OA category, if any, was predictive of clinically relevant improvement. The outcomes of the knee injections and the OA grades were entered into SPSS version 21.0 (Armonk, New York, USA) for analysis. Additionally, the NRS change scores between the patients who had clinically relevant 'improvement' (PGIC 1 and 2) and the patients who did not improve (PGIC 3–7) were assessed for normal data distribution and compared using the unpaired Student's t-test.

► **Table 1** Description of the Kellgren and Lawrence 5 and 3 grades and the Osteoarthritis Research Society International Scoring System [11, 12, 27, 29].

► **Tab. 1** Beschreibungen der Arthrose-Einteilungen nach Kellgren und Lawrence mit 5 und mit 3 Graden sowie des Osteoarthritis Research Society International Einteilungssystems [11, 12, 27, 29].

Kellgren and Lawrence 5-Grade System	
Grade 0	no feature of OA
Grade 1	doubtful narrowing of joint space and possible osteophytic lipping
Grade 2	definite osteophytes, definite narrowing of joint space
Grade 3	moderate multiple osteophytes, definite narrowing of joint space, and some sclerosis and possible deformity of bone ends
Grade 4	large osteophytes, marked narrowing of joint space, severe sclerosis and definite deformity of bone ends
Kellgren and Lawrence 3-Grade System	
Grade 1	no joint space narrowing, no osteophytes, no sclerosis, no cysts, no deformity
Grade 2	definite joint space narrowing, definite osteophytes, slight sclerosis, no cysts, no deformity
Grade 3	gross loss of joint space, definite osteophytes, definite sclerosis, definite cysts, deformity present
Osteoarthritis Research Society International Scoring System	
Grade 0	normal joint space
Grade 1	mild joint space narrowing (1 – 33%)
Grade 2	moderate joint space narrowing (34 – 66%)
Grade 3	severe joint space narrowing (67 – 100%)

Secondary outcome measures compared the NRS change scores (baseline NRS – outcome NRS) between the various KL and OARSI categories using the ANOVA test (parametric data). Pearson's correlation coefficient was used to compare the actual joint space measurements at the 4 different locations to NRS change scores at all follow-up time points. An additional secondary outcome measure included comparing the actual joint space measurements between 'improved' or 'worse' patients using the unpaired Student's t-test, after assessing for normal data distribution.

The inter-observer agreement for the three different grading systems for knee OA was assessed using the Kappa statistic (0 = poor agreement, 0 – 0.20 = slight agreement, 0.21 – 0.40 = fair agreement, 0.41 – 0.60 = moderate agreement, 0.61 – 0.80 = substantial agreement and 0.81 – 1.0 = almost perfect agreement [27]). The inter-rater agreement for the joint space measurements was calculated using the intraclass correlation coefficient (ICC).

## Results

The percentages of patients reporting clinically relevant 'improvement' as well as patients who were unchanged or worse are shown in ► **Table 2**. Looking at the NRS change scores, there

was a statistically significant difference between patients who improved and those who did not improve at one day and one week. At one month the difference was not statistically significant (► **Table 2**).

There were no significant correlations between the actual joint space measurements at any of the 4 measurement sites and the change in the NRS pain scores for any data collection time point.

### Relationship between grading categories and improvement/worsening

There was no statistically significant relationship between the KL3 or KL5 classification systems and 'improvement' (primary outcome) or 'worsening' for any of the data collection time points (p-value range = 0.10 – 0.91 for KL5 and 0.19 – 0.80 for KL3). However, a statistically significant relationship between the OARSI classification and improvement at one day ( $p = 0.004$ ) was found (► **Fig. 4**). Although not statistically significant, there was a tendency for a relationship between KL5 and OARSI with worsening at one week ( $p = 0.095$  and  $p = 0.068$ ). Actual joint space width measurements showed no statistically significant relationship with improvement or worsening after an infiltration at any time point (p-value range = 0.23 – 0.97).

Comparing the NRS change scores based on the KL5, KL3 and OARSI categories showed no significant differences at any time point for either KL classification system. However, once again there was a statistically significant difference ( $p = 0.043$ ) at 1 day for the OARSI classification system. This was due to the difference in the NRS change scores between categories 1 and 2. The mean NRS change score for OARSI grade 1 was 2.40 (SD = 2.55) points and for grade 2 it was 3.66 (SD = 2.41) points.

### Prediction for improvement

To determine which OA grade (OARSI 0 – 3) had the most significant relationship with improvement after an injection, the frequencies of patients who reported clinically relevant 'improvement' within the OARSI groups were calculated. This showed that the proportion of patients reporting clinically relevant improvement within OARSI grade 2 was the highest at all time points, with the highest proportion at one day with 79.4% (► **Fig. 4**). Calculating the odds ratio revealed that patients with OARSI grade 2 are 8 times (8.02) more likely to improve at one day after an infiltration ( $p = 0.024$ ). Patients graded as OARSI 2 were also less likely to report worsening at one day and one month, compared with the other categories, but this was not statistically significant (► **Fig. 4**).

With this model we are able to correctly classify 67.6% of cases and have a positive predictive value of 70.7% and a negative predictive value of 64.2%. The sensitivity of this model is 68.3%, and the specificity is 66.7%.

The proportions of patients reporting 'improvement' or 'worsening' at all data collection time points for the KL5 and KL3 grading systems are also shown in ► **Fig. 5, 6**. There were no significant differences between the different grades at any of the outcome time points for either KL grading system.



► **Fig. 2** **a** Knee with OA grade OARSI (1), KL5 (0), KL3 (1). **b** Knee with OA grade OARSI (2), KL5 (2), KL3 (2). **c** Knee with OA grade OARSI (3), KL5 (4), KL3 (3).

► **Abb. 2** **a** Knie mit Arthrosegrad OARSI (1), KL5 (0), KL3 (1). **b** Knie mit Arthrosegrad OARSI (2), KL5 (2), KL3 (2). **c** Knie mit Arthrosegrad OARSI (3), KL5 (4), KL3 (3).



► **Fig. 3** Knee radiograph with lines indicating locations of joint space measurements.

► **Abb. 3** Knie röntgen mit Linien, welche die Orte der Gelenkspaltmessungen indizieren.

### Inter-observer reliability

The inter-observer reliability showed moderate reliability for grading OA severity with the KL3 and KL5 systems, whereas for the OARSI the reliability was substantial (► **Table 3**). The total agreement within the classification systems was 65 – 76.7%, where the KL5 showed the worst and the OARSI again had the best agreement (► **Table 3**).

► **Table 2** Intra-articular steroid injection outcomes overall.

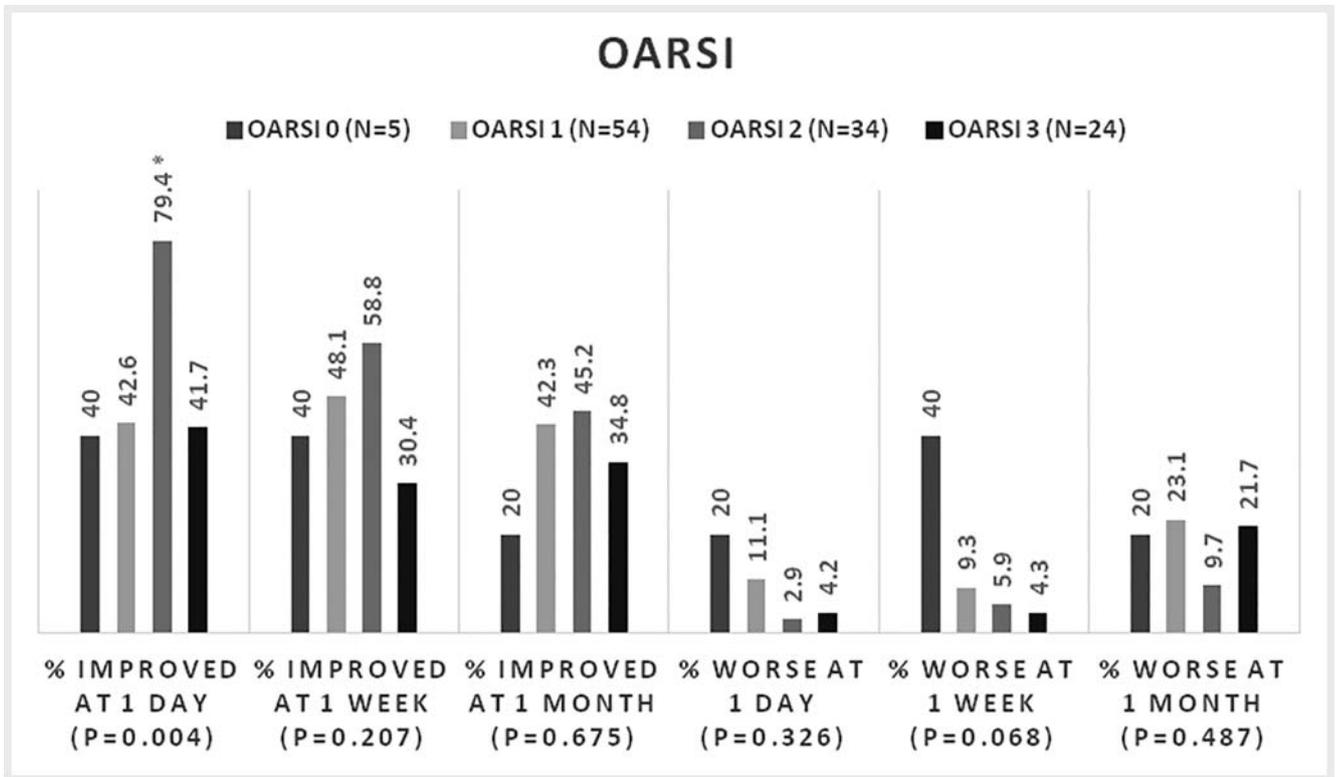
► **Tab. 2** Behandlungsergebnisse von intraartikulären Corticosteroidinjektionen insgesamt.

Proportions of PGIC improved, not improved, worse			
	improved	not improved	worse
1 day	53 %	47 %	7.8 %
1 week	47.4 %	52.6 %	8.6 %
1 month	40.5 %	59.5 %	18.9 %
NRS change scores overall			
	mean NRS change score	SD	
1 day	2.935	2.4861	
1 week	3.091	2.8328	
1 month	2.125	2.8581	
NRS change scores of improved vs. not improved patients, T-test			
	improved	not Improved	p-value
1 day	4.205	1.527	0.001
1 week	3.664	2.444	0.021
1 month	2.263	1.982	0.605

The inter-rater reliability for the actual joint space measurements was very high at all measuring sites (► **Table 3**).

### Discussion

The only OA grading system that had a statistically significant relationship with clinical improvement after intra-articular corticosteroid injections in this study was the OARSI for joint space narrowing. Specifically, a significantly higher percentage of OA



► **Fig. 4** Frequencies of ‘improvement’ or ‘worsening’ by OARSI.

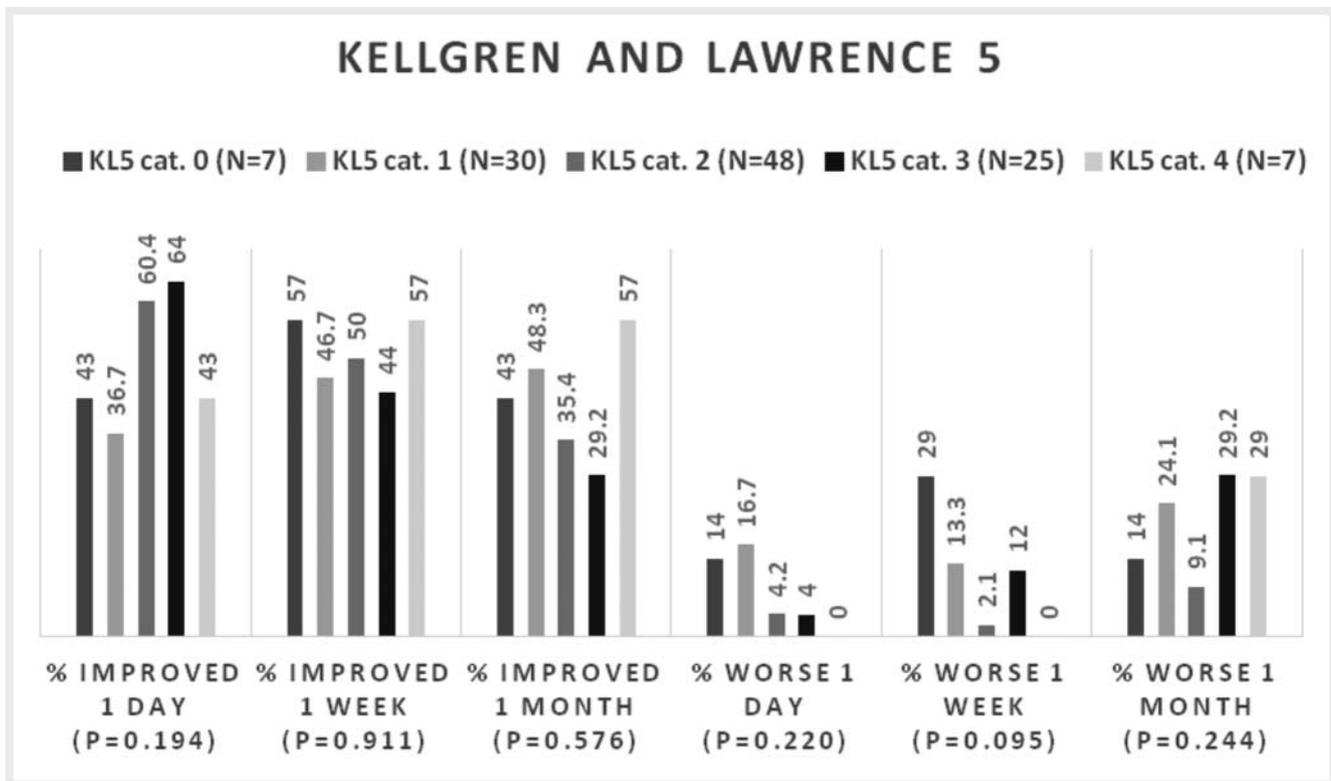
► **Abb. 4** Proportionen von Patienten mit Verbesserung bzw. Verschlechterung innerhalb der Einteilungen nach OARSI.

patients with OARSI grade 2 reported improvement at one day (79.4%) after an infiltration of the knee. Although not statistically significant, the proportions reporting improvement within the OARSI 2 group were also higher at one week and one month and there were also fewer patients reporting worsening at one day and one month compared to the other OARSI grades. The logistic regression analysis supports this finding by showing that there is a significant relationship between the OARSI and improvement at one day. Specifically, patients graded as OARSI 2 are 8 times more likely to report clinically relevant ‘improvement’ after an intra-articular corticosteroid injection compared with the other OARSI grades. Therefore, referring clinicians can have more confidence that their patients graded with OARSI 2 (moderate joint space narrowing of 34–66%) can expect better clinical outcomes, especially at one day after the infiltration compared to the other OARSI grades.

Two systematic reviews were performed in 2013 to try and identify predictors for good treatment response after intra-articular knee injections [28, 29]. Maricar et al. came to the conclusion that the presence of effusion, aspiration of fluid from the knee, severity of disease, absence of synovitis, injection delivery under ultrasound guidance and greater symptoms at baseline may all increase the likelihood of a positive response to intra-articular corticosteroid injections [29]. Hirsch et al., on the other hand, concluded that there is very limited evidence for predictive factors of pain relief following intra-articular corticosteroid injections

in OA of the knee and hip, because the different studies had incongruent results [28]. An interesting fact that might have been related to better outcomes after injection treatments for the knee was a lower radiological degree of degeneration compared to patients with more progressed OA [5, 29, 30]. The results from our current study support these findings, since patients with moderate OA had the best results. To our knowledge, this is the first study to find a possible predictive finding for treatment outcomes in intra-articular corticosteroid injections using the OARSI grading system for joint space narrowing in knee OA, although only with statistical significance at day 1 post-injection. No significant relationships were found when comparing the KL3, KL5 or actual joint space measurements and clinical improvement at any time points.

We chose to compare only the OARSI for joint space narrowing to see if this alone might be a reliable tool to evaluate OA of the knee. We found that not only was the OARSI for joint space narrowing the only grading system linked with significant clinical improvement, but it also had the best inter-observer reliability compared to the KL3 and KL5 grading systems. While the inter-observer agreement for the OARSI was substantial, it only reached moderate agreement for the KL3 and KL5. Only the actual joint space measurements reached higher and almost perfect inter-rater agreement. However, since the OARSI grading system for joint space narrowing relies only on the proportion of joint space width (divided into thirds), the fact that this grading system cor-



► Fig. 5 Frequencies of 'improvement' or 'worsening' by KL5.

► Abb. 5 Proportionen von Patienten mit Verbesserung bzw. Verschlechterung innerhalb der Einteilungen nach KL5.

responds well with the results of the actual joint space measurements is not surprising. The OARSI system was also very quick and easy to use.

An earlier study suggested that the KL scales may require a re-appraisal [20]. The most notable reasons were inconsistencies in the descriptions of radiographic features of osteoarthritis by Kellgren and Lawrence themselves as well as in other studies [20, 25], the prominence awarded to the osteophytes at all joint sites [10, 13, 14, 17, 20] and the relative insensitivity to change [14, 17]. While in fact osteophytes are most strongly associated with knee pain [17, 19, 20], OA progression does not follow a specific pattern that always starts with osteophytes. The typical OA cut-off using the KL system, however, is  $\geq$  grade 2, while the cut-off for the OARSI atlas entails meeting any one of three separate criteria: either joint space narrowing grade  $\geq 2$ , sum of osteophyte grades  $\geq 2$  or grade 1 joint space narrowing in combination with a grade 1 osteophyte. Therefore, to diagnose knee OA with the KL system, definitive osteophytes need to be present. Additionally, unlike the KL system, the OARSI atlas grades features in the medial and lateral femorotibial compartments separately, which helps to capture those with early joint changes more effectively [8]. When performing the radiographic readouts, it was our experience that the step from grade 2 to grade 3 in the KL5 system was the most problematic with the lowest level of agreement. The decision was often made on the basis of impression rather than facts, since OA progression does not follow a strict pattern in all patients and there was considerable overlap in

the radiographic findings particularly between these 2 categories. Schiphof et al. reported on how the original KL system was adapted in different studies and found that most adaptations were made with grade 2 [25]. Riddle et al. also support our difficulty of rating lower grades of OA by finding that the inter-observer agreement was better for more progressed OA [24]. With the KL3 grading system, it was slightly more difficult to differentiate grade 1 (no signs of OA) from grade 2 (definite osteophytes, definite joint space narrowing and slight sclerosis), because the step was a little too strict and several patients would have joint space narrowing but no osteophytes or vice versa. Because the OARSI evaluation method for joint space narrowing alone is easy and a more objective and reliable tool for evaluating the severity of knee OA, it can easily be implemented in clinical practice, especially if it can be used as a predictor of a positive response to intra-articular steroid injections.

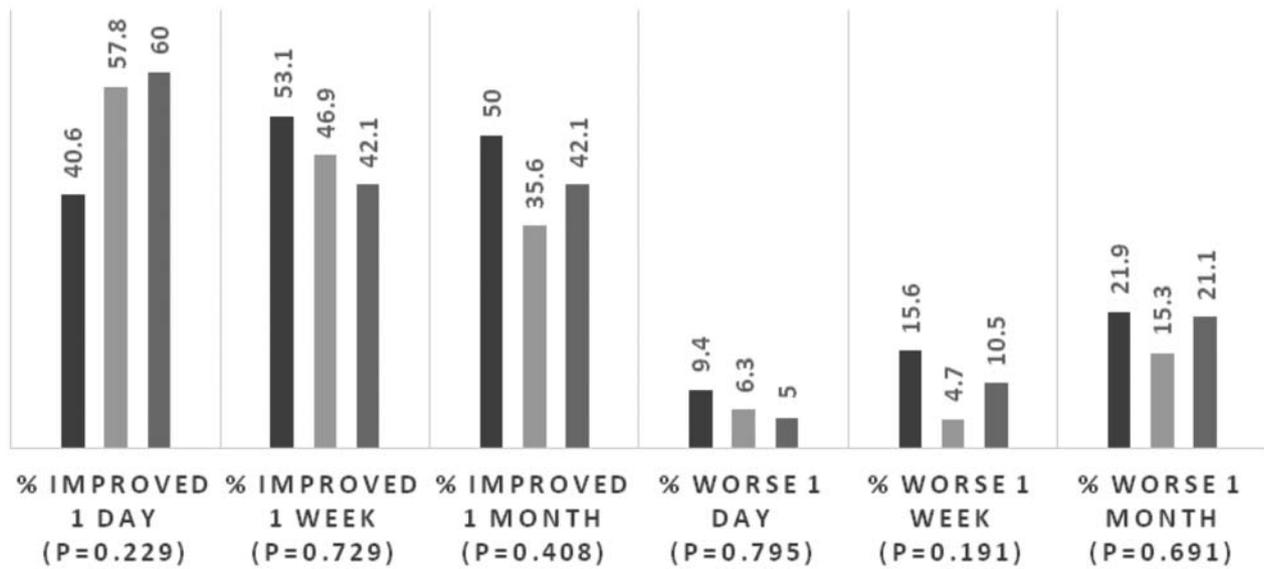
### Limitations

One limiting factor of this study is the distribution of severity of OA grades, which was not equal for the different categories. Especially OARSI grade 0 was limited, with only five patients. Perhaps a larger sample size with better distribution among the various OA severity categories would provide clearer results, particularly for comparing patients in OARSI categories 0 and 2.

The patient follow-up of one month might seem rather short. However, since earlier studies only support clinically relevant improvement from intra-articular corticosteroid injections for up

## KELLGREN AND LAWRENCE 3

■ KL3 cat. 1 (N=32) ■ KL3 cat. 2 (N=64) ■ KL3 cat. 3 (N=20)



► **Fig. 6** Frequencies of 'improvement' or 'worsening' by KL3.

► **Abb. 6** Proportionen von Patienten mit Verbesserung bzw. Verschlechterung innerhalb der Einteilungen nach KL3.

► **Table 3** Reliability of Classification Systems and Measurements.

► **Tab. 3** Reliabilität der Einteilungssysteme sowie der effektiven Messungen.

interobserver reliability kappa/agreement

KL3	0.554 = moderate	76.5 %
KL5	0.482 = moderate	65.0 %
OARSI	0.660 = substantial	76.7 %

ICC for joint space measurements

site of measurement	ICC value	95 % confidence interval
middle of medial plateau	0.882 = almost perfect	0.834 – 0.917
middle of lateral plateau	0.812 = almost perfect	0.739 – 0.866
edge of medial plateau	0.876 = almost perfect	0.826 – 0.912
edge of lateral plateau	0.812 = almost perfect	0.738 – 0.867

KL = Kellgren and Lawrence; OARSI = Osteoarthritis Research Society International; ICC = Intraclass correlation coefficient.

to three weeks [1 – 5], one month should be long enough to evaluate clinical outcomes and compare them with OA grades.

The outcomes after intra-articular corticosteroid injections were collected in a prospective manner in this study. The evalua-

tion of the different OA grades was performed retrospectively. Even though the evaluation was done blinded to the treatment outcomes, the retrospective character might be considered as a limitation.

## Conclusion

Comparing the Kellgren and Lawrence 5- and 3-grade systems, the Osteoarthritis Research Society International Grading System (OARSI) for joint space narrowing and actual joint space measurements, we found that not only is the OARSI system the only one that was predictive of better outcomes after an intra-articular corticosteroid injection at 1 day, but it also had better inter-observer reliability than the KL5 and KL3 grading systems. Especially patients with OARSI grade 2 seem to report significantly more improvement at one day and, although not statistically significant but possibly clinically relevant, better improvement at one week and one month as well. To our knowledge, this is the first study to find a possible predictive value for intra-articular corticosteroid injections using the OARSI grading system for joint space narrowing in knee OA. This could be an easy and more reliable tool in clinical practice to predict which patients could benefit most from intra-articular corticosteroid injections.

### CLINICAL RELEVANCE OF THE STUDY

- The OARSI grading of OA had better reliability than either the KL3 or KL5 grading system.
- OARSI grade 2 was related to better treatment outcomes.
- Neither the KL grades nor the joint space measurements were related to improvement.

## Conflict of Interest

The authors declare that they have no conflict of interest.

## References

- [1] Bellamy N, Campbell J, Robinson V et al. Intraarticular corticosteroid for treatment of osteoarthritis of the knee. *Cochrane Database Syst Rev* 2006; 2: CD005328
- [2] Gaffney K, Ledingham J, Perry JD. Intra-articular triamcinolone hexacetonide in knee osteoarthritis: factors influencing the clinical response. *Ann Rheum Dis* 1995; 54: 379–381
- [3] Godwin M, Dawes M. Intra-articular steroid injections for painful knees. Systematic review with meta-analysis. *Can Fam Physician* 2004; 50: 241–248
- [4] Jones A, Doherty M. Intra-articular corticosteroids are effective in osteoarthritis but there are no clinical predictors of response. *Ann Rheum Dis* 1996; 55: 829–832
- [5] Smith MD, Wetherall M, Darby T et al. A randomized placebo-controlled trial of arthroscopic lavage versus lavage plus intra-articular corticosteroids in the management of symptomatic osteoarthritis of the knee. *Rheumatology* 2003; 42: 1477–1485
- [6] Raynauld JP, Buckland-Wright C, Ward R et al. Safety and efficacy of long-term intraarticular steroid injections in osteoarthritis of the knee: a randomized, double-blind, placebo-controlled trial. *Arthritis Rheum* 2003; 48: 370–377
- [7] Jüni P, Reichenbach S, Trelle S et al. Efficacy and safety of intraarticular hylan or hyaluronic acids for osteoarthritis of the knee: a randomized controlled trial. *Arthritis Rheum* 2007; 56: 3610–3619
- [8] Culvenor AG, Engen CN, Øiestad BE et al. Defining the presence of radiographic knee osteoarthritis: a comparison between the Kellgren and Lawrence system and OARSI atlas criteria. *Knee Surg Sports Traumatol Arthrosc* 2015; 23: 3532–3539
- [9] Petersson IF, Boegård T, Saxne T et al. Radiographic osteoarthritis of the knee classified by the Ahlbäck and Kellgren & Lawrence systems for the tibiofemoral joint in people aged 35–54 years with chronic knee pain. *Ann Rheum Dis* 1997; 56: 493–496
- [10] Hart DJ, Spector TD. The classification and assessment of osteoarthritis. *Baillieres Clin Rheumatol* 1995; 9: 407–432
- [11] Blackburn WD, Bernreuter WK, Rominger M et al. Arthroscopic evaluation of knee articular cartilage: a comparison with plain radiographs and magnetic resonance imaging. *J Rheumatol* 1994; 21: 675–679
- [12] Brandt KD, Fife RS, Braunstein EM et al. Radiographic grading of the severity of knee osteoarthritis: relation of the Kellgren and Lawrence grade to a grade based on joint space narrowing, and correlation with arthroscopic evidence of articular cartilage degeneration. *Arthritis Rheum* 1991; 34: 1381–1386
- [13] Kijowski R, Blankenbaker D, Stanton P et al. Arthroscopic validation of radiographic grading scales of osteoarthritis of the tibiofemoral joint. *Am J Roentgenol* 2006; 187: 794–799
- [14] Altman RD, Gold GE. Atlas of individual radiographic features in osteoarthritis, revised. *Osteoarthritis Cartilage* 2007; 15: A1–A56
- [15] Eckstein F, Wirth W, Hunter DJ et al. Magnitude and regional distribution of cartilage loss associated with grades of joint space narrowing in radiographic osteoarthritis—data from the Osteoarthritis Initiative (OAI). *Osteoarthritis Cartilage* 2010; 18: 760–768
- [16] Kellgren JH, Lawrence JS. Radiological assessment of osteoarthrosis. *Ann Rheum Dis* 1957; 16: 494–502
- [17] Altman R, Brandt K, Hochberg M et al. Design and conduct of clinical trials in patients with osteoarthritis: recommendations from a task force of the Osteoarthritis Research Society. Results from a workshop. *Osteoarthritis Cartilage* 1996; 4: 217–243
- [18] Altman RD, Fries JF, Bloch DA et al. Radiographic assessment of progression in osteoarthritis. *Arthritis Rheum* 1987; 30: 1214–1225
- [19] Hellio Le Graverand MP, Mazzuca S, Duryea J et al. Radiographic-based grading methods and radiographic measurement of joint space width in osteoarthritis. *Radiol Clin North Am* 2009; 47: 567–579
- [20] Spector TD, Cooper C. Radiographic assessment of osteoarthritis in population studies: whither Kellgren and Lawrence? *Osteoarthritis Cartilage* 1993; 1: 203–206
- [21] Humphreys BK, Peterson C. Comparison of outcomes in neck pain patients with and without dizziness undergoing chiropractic treatment: a prospective cohort study with 6 month follow-up. *Chiropr Man Therap* 2013; 21: 3
- [22] Peterson CK, Bolton J, Humphreys BK. Predictors of improvement in patients with acute and chronic low back pain undergoing chiropractic treatment. *J Manipulative Physiol Ther* 2012; 35: 525–533
- [23] Peterson C, Bolton J, Humphreys BK. Predictors of outcome in neck pain patients undergoing chiropractic care: comparison of acute and chronic patients. *Chiropr Man Therap* 2012; 20: 27
- [24] Riddle DL, Jiranek WA, Hull JR. Validity and reliability of radiographic knee osteoarthritis measures by arthroplasty surgeons. *Orthopedics* 2013; 36: e25–e32
- [25] Schiphof D, Boers M, Bierma-Zeinstra SM. Differences in descriptions of Kellgren and Lawrence grades of knee osteoarthritis. *Ann Rheum Dis* 2008; 67: 1034–1036
- [26] Terjesen T, Gunderson RB. Radiographic evaluation of osteoarthritis of the hip: an inter-observer study of 61 hips treated for late-detected developmental hip dislocation. *Acta Orthop* 2012; 83: 185–189

- [27] Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33: 159–174
- [28] Hirsch G, Kitas G, Klocke R. Intra-articular corticosteroid injection in osteoarthritis of the knee and hip: factors predicting pain relief—a systematic review. *Semin Arthritis Rheum* 2013; 42: 451–473
- [29] Maricar N, Callaghan MJ, Felson DT et al. Predictors of response to intra-articular steroid injections in knee osteoarthritis—a systematic review. *Rheumatology* 2013; 52: 1022–1032
- [30] Evanich JD, Evanich CJ, Wright MB et al. Efficacy of intraarticular hyaluronic acid injections in knee osteoarthritis. *Clin Orthop Relat Res* 2001; 390: 173–181