European Journal of Radiology 88 (2017) 95–101

Contents lists available at ScienceDirect

European Journal of Radiology

journal homepage: www.elsevier.com/locate/ejrad

Multirater agreement for grading the femoral and tibial cartilage surface lesions at CT arthrography and analysis of causes of disagreement☆

Patrick Omoumi [a,b,*], Nicolas Michoux [a], Ahmed Larbi [a], Laure Lacoste [a], Frédéric E. Lecouvet [a], Vasiliki Perlepe [a], Bruno C. Vande Berg [a]

[a] Department of Radiology, Cliniques Universitaires St Luc – UC Louvain, Hippocrate Avenue 10/2942, B-1200 Brussels, Belgium
[b] Department of Diagnostic and Interventional Radiology, Lausanne University Hospital, Bugnon 46, CH-1011 Lausanne, Switzerland

ARTICLE INFO

Article history:
Received 13 October 2016
Received in revised form
21 November 2016
Accepted 24 December 2016

Keywords:
Cartilage
Osteoarthritis
Thickness
Grading system
Computed tomography
Arthrography

ABSTRACT

Objective: To assess the multirater agreement of the modified Outerbridge system for the grading of pre-defined areas of femorotibial cartilage at CT arthrography with multiple readers, with varying experience.
Design: Five readers with varying experience (two junior radiologists, three musculoskeletal radiologists including two experts in cartilage imaging) separately analyzed 962 cartilage sectors from pre-divided knee CT arthrograms with femorotibial osteoarthritis (Kellgren/Lawrence = 3). Each cartilage area was graded twice by each reader, at a three-month interval, according to the modified 5-grade Outerbridge system. Interobserver and intraobserver agreement were assessed. After the second reading, 121 areas exhibiting the highest interobserver disagreement were reviewed in consensus to determine the sources of disagreement.
Results: The global interobserver agreement was fair (k = 0.35), and increased with the grade (from k = 0.14 to k = 0.76 from grade 0–4). The intraobserver agreement varied with the readers' experience from moderate (k = 0.59) to almost perfect (k = 0.92). The majority of cases of disagreement (44%) was due to difficulties in assessing the normal variations of cartilage thickness, including diffuse cartilage thinning (23%) and normal variants of cartilage thickness (22%). 32% of cases of disagreement were due to retrospectively avoidable interpretation errors.
Conclusions: The multirater agreement of the modified Outerbridge system is only fair when readers of different level of experience are taken into account, and interobserver agreement increases with readers' experience. However, interobserver agreement is substantial for grade 4 lesions. We report normal variations of cartilage thickness that may improve observer agreement in reporting cartilage lesions.

Introduction

Cartilage plays a fundamental role in the biomechanics of joints, both by decreasing friction at joint surfaces and by absorbing and distributing load across the joint [1].

Cartilage has limited intrinsic repair capacities, which has several consequences. First, cartilage lesions, usually due to traumatic or degenerative cause, seldom heal and may progress to osteoarthritis, causing pain along the course of the disease [1,2]. Second, treating cartilage lesions has proven to be quite challenging, and despite years of research, there is still no efficient treatment to slow or stop the progression of cartilage lesions towards osteoarthritis [3,4].

One important factor to consider in the management of cartilage lesions is the depth of the cartilage lesions [1,3,5]. Therefore, the assessment of the depth of cartilage substance loss relative to the normal cartilage has been the basis for all grading systems [6–11]. These systems have been widely used for years, both in clinical practice to routinely report the severity of chondral abnormalities at cross-sectional imaging, as well as in research settings, to develop and validate new therapeutic solutions.

The standard grading system for the evaluation of cartilage lesions is the Outerbridge system, which was first described at

☆ Department to which the work should be attributed: Department of Radiology, Cliniques Universitaires St Luc - UC Louvain, Hippocrate Avenue 10/2942, B-1200 Brussels, Belgium.

* Corresponding author at: Department of Diagnostic and Interventional Radiology, Lausanne University Hospital, Bugnon 46, CH-1011 Lausanne, Switzerland.
E-mail address: patrick.omoumi@chuv.ch (P. Omoumi).

http://dx.doi.org/10.1016/j.ejrad.2016.12.026
0720-048X/

arthroscopy for chondromalacia patella and later adapted by Noyes to the entire knee joint [12,13]. Later, it was applied to other joints and adapted to cross-sectional imaging, both computed tomography (CT) arthrography and magnetic resonance imaging (MRI), to enable non-invasive grading of cartilage lesions [6,7]. Subsequent grading systems for the assessment of cartilage have been developed and are based on the same principle (evaluation of the depth of substance loss), with the adjunct of other imaging features such as the size of the lesions, and intrasubstance chondral signal abnormalities at MRI [8–11].

Several studies have previously shown that the interobserver agreement of the modified Outerbridge classification at arthroscopy is limited, mostly fair or moderate, and it depends on the surgeons' experience [14–16]. In practice, the application of such grading systems based on the depth of the lesions can also be difficult at imaging. To the best of our knowledge, the modified Outerbridge grading system has never been assessed on more than two readers with varying levels of experience.

In this study, we sought to evaluate the inter- and intra-observer agreement of the widely used five-grade modified Outerbridge system with five readers of varying level of experience, using CT arthrography. We chose CT arthrography as the cross-sectional modality presenting the highest contrast and spatial resolution for the study of cartilage, allowing the easiest evaluation of cartilage surface lesions [17,18]. In particular, the great attenuation difference between cartilage and the articular iodinated contrast material filling the lesions allows easier detection of cartilage substance loss than with MRI, and higher diagnostic confidence [18,19]. We also aimed to analyze sources of disagreement between readers to determine potential pitfalls in reporting cartilage lesions.

2. Method

The following paragraphs detail the methodology that we developed to achieve the following goals: 1. to assess the multirater agreement with readers of varying experience, 2. on a large number of cartilage areas, 3. while ensuring that the exact same areas would be graded by all readers.

2.1. Patient population

Knee CT arthrograms from five consecutive patients (3 females, mean age: 78 [55–84]) with Kellgren-Lawrence grade = 3 femorotibial osteoarthritis at radiography were retrospectively included. We excluded any examination presenting motion artifacts, or an inhomogeneous coverage of articular cartilage surface by contrast material. This study was approved by our institutional ethical committee and no patient consent was required due to its retrospective nature.

2.2. Imaging technique

Arthrography was performed by injecting a volume of 10 ml of ionic contrast material [meglumine ioxalate and sodium ioxalate; Hexabrix 320 (320 mg of iodine per milliliter); Guerbet, Aulnay-sous-bois, France] into the knee under fluoroscopic guidance. CT examinations were performed on a 40-slice multidetector CT (Somatom Definition AS; Siemens Healthcare, Forchheim, Germany) within 15 min after the injection, to avoid significant penetration of contrast media into cartilage substance [20]. All patients were lying supine with the knee in extension. Acquisition parameters were optimized for the knee joint: 120 kVp and 350 mAs with the application of a dose modulation protocol (Care Dose 4D; Siemens Healthcare); detector configuration: 16×0.6 mm; pitch: 0.85; gantry rotation time: 1 s [21]. Image reconstruction parameters were: field-of-view (FOV): 15×15 cm; matrix: 512^2; section thickness/increment: 0.6/0.3 mm; bone convolution kernel (U70u). All examinations were stored on our PACS system (Carestream Client version 11.3; Carestream Health, Rochester, NY, USA).

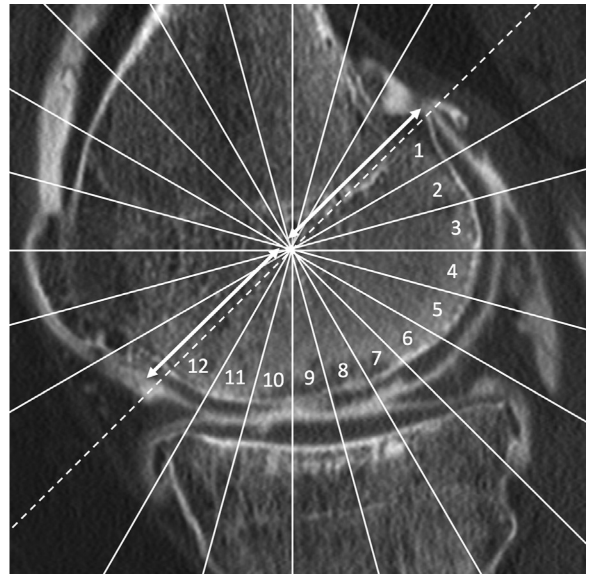

Fig. 1. Midsagittal reformat of CT arthrogram of a medial knee compartment showing anteroposterior segmentation of femorotibial cartilage. A grid with 24 radial segments of 15° each was pasted on each sagittal reformat. The grid was panned so that 1/the −45° line (dashed line) would intersect the most posterior aspect of the distal femoral physeal line and 2/the center of the grid would project halfway between the two intersections of this −45° line and the condylar cartilage (arrows of equal length). On this slide, 12 cartilage segments, as indicated by the numbers, were analyzed by each of the five readers.

2.3. Image preparation

Prior to the readings, each femorotibial cartilage surface was divided into small areas, using a method that ensured that all readers analyzed the same cartilage areas on these pre-divided images. One observer who was not part of the readers processed the examinations. The images were zoomed in by a factor of 2.5 so that they would optimally fill the available screen space. Each femorotibial compartment was divided into ten 1mm-thick reformats in the sagittal plane, perpendicular to a line joining the most posterior aspect of femoral condyles. Each sagittal reformat was exported in TIFF format into PowerPoint slides (maximizing the image size to fill the slide). Each femorotibial compartment was then segmented anteroposteriorly into up to 15 areas using the method described in Fig. 1 [22]. A grid with 24 radial segments of 15° each was drawn in PowerPoint, and copy pasted on the midsagittal reformat of each condyle. The grid was then moved on that image to fill two conditions: 1/the −45° line (dashed line in Fig. 1) should intersect the most posterior aspect of the distal femoral physeal line and 2/the center of the grid should project halfway between the two intersections of this −45° line and the condylar cartilage (arrows of equal length in Fig. 1). Once properly placed on the midsagittal plane of each condyle, the grid was then copy-pasted on all sagittal formats of the same condyle. The same pre-divided images were given to each observer for the readings (Fig. 2).

2.4. Readings

Readings were performed on a 15-inch computer screen with a resolution of 1920 by 1200 pixels, analyzing the slides in full screen mode.