AUTOMATED KNEE OSTEOARTHRITIS ASSESSMENT INCREASES PHYSICIANS’ AGREEMENT RATE AND ACCURACY: DATA FROM THE OSTEOARTHRITIS INITIATIVE
Paixao T., Stria A., Goetz C., Ljuhar R., Bertalan Z.

Objective: The diagnosis of knee osteoarthritis depends on the identification and classification of several radiographic features, such as presence and degree of osteophytes, sclerosis, and joint space narrowing. Here, we assess the impact of a computerized system on physicians’ accuracy and agreement rate, as compared to unaided diagnosis.

Methods: A set of 124 unilateral knee radiographs from the OAI study were selected and analyzed by a deep learning-based method with regard to Kellgren-Lawrence (KL) grade, as well as Joint Space Narrowing, Osteophytes and Sclerosis OARSI grades. Physicians were instructed to score all images, with respect to these features, in two modalities: being shown simply the image of a radiograph (unaided) and when presented with the report from the computer assisted detection system (aided). The two reading sessions were separated by an appropriate washout period. The readers were blinded to each other’s grades and to the ground truth grading (OAI consensus grades). Agreement rates (Intra-Class Correlation - ICC) between the physicians were calculated for both modalities. Furthermore, the physicians’ performance was compared to the ground truth grading (OAI consensus), and accuracy, sensitivity and specificity in both modalities were calculated for each feature.

Results: Agreement rates (ICC) for KL grade, sclerosis, and osteophyte OARSI grades, were statistically increased in the aided modality vs the unaided modality. Readings for Joint Space Narrowing OARSI grade did not show a statistical difference between the two modalities. Readers’ accuracy for detection of any abnormality (KL>0), osteoarthritis (KL>1), sclerosis (sclerosis OARSI grade > 0), and osteophytosis (osteophyte OARSI grade > 0) was significantly increased in the aided modality. These increases in accuracy were driven by significant increases in specificity, with no statistical difference in sensitivity.

Conclusions: These results show the use of an automated knee osteoarthritis software increases consistency between physicians when grading radiographic features of OA. Furthermore, the use of a software solution increases specificity with no losses in sensitivity.
Objective: The rate of cartilage loss can vary wildly between patients at risk or suffering or from knee osteoarthritis (OA) but its causes remain unknown. We investigate whether quantitative and semi-quantitative radiographic features can be used to predict the rate of Joint Space Width (JSW) loss.

Methods: We collected bilateral knee radiographs, acquired in the context of the OAI study, from 4100 patients (2383 female, 1717 male). Each patient was imaged up to 7 times, separated by at least 12 months, across a time span of 8 years. Each radiograph was analyzed by software to obtain Kellgren-Lawrence (KL) and OARSI grades for Osteophytes, Sclerosis and Joint Space Narrowing (JSN) readings, as well as JSW measurements for each individual knee. Linear regressions of JSW were performed per individual knee compartment (medial or lateral) to estimate the rate of JSW loss per month. Individual knees with rate of JSW loss above 0.072 mm/year (the average yearly loss within JSN grade) were classified as progressors (956 knees). From these, knees in the top 10% of JSW loss rate were classified as fast progressors (91 knees). A logistic regression model was used to predict the fast progressor phenotype with KL and OARSI grades at baseline as independent variables. Model performance was estimated using 10-fold cross-validation training/testing dataset splits and used Area Under the Curve (AUC) as performance criteria.

Results: The logistic regression classifiers achieved AUCs of 0.71 (SE 0.015) and 0.66 (SE 0.013) at classifying individual knees as fast progressors for medial and lateral compartments, respectively. Analysis of the individual coefficients of the classifiers reveals that JSN and Sclerosis OARSI grades are the main predictors of rapid cartilage loss.

Conclusion: Our results show that it is possible to predict future rapid cartilage loss from quantitative and semi-quantitative readings from a single plain radiograph. Interestingly, neither KL grade nor Osteophytes OARSI grade contributed greatly to this prediction. Instead, Sclerosis and JSN grade seem to be the major predictors of rapid cartilage loss, suggesting a non-canonical mode of OA progression.
A NOVEL QUANTITATIVE METRIC FOR JOINT SPACE WIDTH: DATA FROM THE OSTEOARTHRITIS INITIATIVE (OAI)

Paisao T., Goetz C., Bertalan Z., Ljuhar D., Ljuhar R., Nehrer S.

Purpose: Joint Space Width (JSW) has been the gold standard to assess loss of cartilage in knee OA. Here we describe a novel quantitative measure of joint space width, standardized JSW (stdJSW). We assess the performance of this quantitative metric for joint space width that tracking Joint Space Narrowing OARSI grade (JSN) changes and provide reference values for different joint space narrowing OARSI grades and their annual change.

Methods: We collected 18,934 individual knee images from the OAI study, from the follow-up visits up to month 48 (baseline plus 4 follow-up exams). Absolute JSW measurements and JSN readings were collected from the OAI study. Standardized JSW was calculated for each knee as well as 12-month JSN grade changes. For each JSN grade and 12-month grade change, the distribution of JSW loss was calculated both for standardized JSW as well as absolute JSW measurements retrieved from the OAI study. Area under the curve of the ROC curves was calculated for the performance of both absolute and standardized JSW at discriminating between different JSN grades. Standardized response mean (SRM) was used to compare the responsiveness of the two measures to change in JSN grade.

Results: The areas under the ROC curve for stdJSW at discriminating between successive JSN grades were AUC_{stdJSW} = 0.87, 0.95, and 0.96, for JSN>0, JSN>1 and JSN>2, respectively, whereas these were AUC_{JSW} = 0.79, 0.90, 0.98 for absolute JSW. We find that standardized JSW is significantly more responsive than absolute JSW, as measured by the Standardized Response Mean. Furthermore, we present reference values for standardized JSW stratified by base JSN grade and 12-month JSN change.

Conclusions: Our results show that standardized JSW is a better choice to track changes in JSN and to discriminate between JSN grades. Furthermore, our results show that standardized JSW performs better in part because it cancels part of the normal variation in JSWs that comes from variation in height.
PREDICTING JOINT NARROWING RISK FROM A NOVEL JOINT SPACE WIDTH MEASURE: DATA FROM THE OSTEOARTHRITIS INITIATIVE
Paixao T., Goetz C., Bertalan Z., Ljuhar D., Ljuhar R., Nehrer S.

ABSTRACT

Objective: To assess the ability of a novel joint space width measure to determine risk of OA progression.

Methods: 4100 individual knee radiographs were selected from the OAI study. Joint space narrowing OARSI grades (JSN) were collected for the visits corresponding to the first 48 months. Individual knees were classified as progressors or not, based on the JSN scores at every 12-month interval. The selected radiographs were subsequently analyzed by a computer assisted detection system (IB Lab’s KOALA) and the measurements of joint space width were retrieved. Logistic regression models were trained separately for the medial and lateral compartments to classify knees as progressors or non-progressors, based solely on these JSW measurements of 1, 2 or 3 consecutive 12 month assessments. Model performance was assessed based on the Area under the ROC curve (AUC) in a 10-fold stratified splitting of the data.

Results: Models that used only JSW measurements for one visit achieved a performance of 0.88 (0.76-1.00) and 0.73 (0.59-0.87) for the lateral and medial compartments, respectively. Models that made use of two sets of measurements separated by 12 months, achieved a performance of 0.92 (0.82-1.00) and 0.79 (0.57-1.00), for lateral and medial compartments, respectively. Models using 3 consecutive sets of JSW measurements achieved performances of 0.90 (0.65-1.0) and 0.81 (0.65-0.97) for the lateral and medial side, respectively.

Conclusions: Our results show that a novel measure of the joint space width can predict the risk of OA progression. Models trained for the lateral compartments achieved excellent performance when using data for the two visits separated by a 12 month interval. Models for the medial compartment require at least 3 consecutive sets of measurements in order to achieve good performance.