15th International Workshop on Osteoarthritis Imaging

Rotterdam, Netherlands

June 30 – July 2, 2021

WELKOM!

Dear IWOAI attendee,

We extend a warm welcome to Rotterdam to all of you, regardless of whether you are joining the 15th International Workshop on Osteoarthritis Imaging on-line or on the ground in our beautiful city. We hope you are as excited as we are about this hybrid meeting which, despite the continuing COVID-19 crisis in many parts of the world, continues an established annual tradition to bring together scientists, clinicians and industrial partners with a heart for imaging of osteoarthritis.

Those of you who have overcome travel restrictions and are joining us in person will have a chance to experience Rotterdam as a vibrant port city, where exceptional art meets revolutionary architecture as well as a fantastic culinary and nightlife scene. You will without doubt appreciate Rotterdam as a city of innovation and understand why it ranks as a top European tourist destination. IWOAI 2021 is brought to you from our venue located along the river Maas and directly adjacent to the iconic Erasmus Bridge, and boasts a magnificent view of the Rotterdam skyline. We hope that this will allow remote attendees to also catch an impression of our city.

The theme of this year's meeting is "Open Up: The Multifaceted Nature of OA Imaging". It emphasizes the purpose of the IWOAI to be inclusive and to unite researchers from various disciplines, each with their own perspectives on OA imaging which will be highlighted in a panel discussion. Related to this, another panel discussion will address multiple perspectives on the open data paradigm applied to OA imaging research. We have put every effort into creating an attractive and inspiring program for you, which features eminent speakers and cutting-edge research covering a wide range of topics linked to OA imaging, and we hope you will enjoy it.

The unique strength of IWOAI lies in the combination high-level scientific content with ample opportunities for social interaction, networking, and fun, which has led to productive collaborations and long-standing friendships. Unfortunately, we cannot have a social program in the usual format, but we have found an exciting and contemporary alternative by means of a virtual social environment on Gather.Town. This virtual platform will be open throughout the meeting and will serve as our base to virtually gather outside of the plenary sessions. Furthermore, we will hold our poster sessions in the virtual poster hall on Gather.Town, providing the opportunity for interaction in a similar fashion as traditional poster sessions.

Thank you for being part of IWOAI 2021 and enjoy the meeting!


Edwin Oei and Jos Runhaar
IWOAI 2021 co-Chairs

15th International Workshop on Osteoarthritis Imaging
Open Up: The Multifaceted Nature of OA Imaging

Chairs

Edwin Oei, MD, PhD (Erasmus MC, dept. Radiology & Nuclear Medicine)
Jos Runhaar, PhD (Erasmus MC, dept. General Practice)

Local Organizers






- Sita Bierma-Zeinstra, PhD (Erasmus MC, Depts. of General Practice and Orthopedics & Sports Medicine)
- Susanne Eijgenraam, MD, PhD (Erasmus MC, Depts. Radiology & Nuclear Medicine)
- Rianne van der Heijden, MD, PhD (Erasmus MC, Depts. Radiology & Nuclear Medicine)
- Jukka Hirvasniemi, PhD (Erasmus MC, Depts. Radiology & Nuclear Medicine)
- Stefan Klein, PhD (Erasmus MC, Depts. Radiology & Nuclear Medicine)
- Dieuwke Schiphof, PhD (Erasmus MC, Dept. General Practice)

Scientific Committee

- Shadpour Demehri, MD
- John Lynch, PhD
- Emily McWalter, PhD
- Sven Nebelung, MD
- Valentina Pedoia, PhD
- Tom Turmezei, MD
- Wolfgang Wirth, PhD

Wednesday June 30, 2021

All times are Rotterdam time (CET)

13.00 – 14.00	Lunch at Inntel
14.15 – 14.30	**Opening:** Edwin Oei & Jos Runhaar (Chairs of the meeting)
14.30 - 16.00	**Session 1: "Imaging of early OA; patient selection vs. outcome measures"**
14.30 – 15.00	Keynote speaker: Sita Bierma-Zeinstra Moderators: Simo Saarakkala & Wolfgang Wirth
15.00 – 16.00	Four Abstracts Roemer: Kellgren and Laurence grade 2 and 3 knees exhibit heterogeneous spectrum of cartilage damage: The MOST study Wisser: Association of semiquantitative MRI-based measures of knee cartilage with increased cartilage loss in knees at elevated risk of developing OA – Data from the Osteoarthritis Initiative Runhaar: Evaluating the tibial spines as imaging biomarker for incident knee OA Panfilov: Deep learning to predict early radiographic knee OA progression directly from MRI
16.00 – 17.00	Break & social gathering on Gather.Town
17.00 – 18.30	**Session 2: "New (MRI) techniques for OA imaging"**
17.00 – 17.30	Keynote speaker: Patrick Omoumi Moderators: Jutta Ellerman & Feliks Kogan
17.30 – 18.30	Four abstracts Heiss: Compositional MRI of articular cartilage of the wrist at 3T and 7T MRI: A comparative study of T2 and T2* relaxometry Einarsson: Pilot imaging ex vivo human meniscus with phase-contrast enhanced synchrotron micro-tomography Kestila: 3D grading of calcifications from ex vivo human meniscus posterior horn Egnell: Evaluation of an AI system for knee osteoarthritis
18.30 – 18.40	**Acknowledgements of Young Investigator Award winners** By: Emily McWalter & John Lynch Jansen: Knee joint distraction results in MRI cartilage thickness increase up to ten years after treatment Anwari: Knee pain is associated with perfusion kinetics at the infrapatellar fat pad among non-overweight postmenopausal women Nguyen: Interpretable deep learning framework for knee osteoarthritis structural prognosis prediction
18.40 – 19.00	**IWOAI/ISOAI businesses** Ali Guermazi, ISOAI President Frank W. Roemer, Editor in Chief, Osteoarthritis Imaging

Thursday July 1, 2020

All times are Rotterdam time (CET)

08.30 – 10.30	Local social outing (pending COVID regulations)
10.30 – 11.00	Pre-meeting update on Computed Tomography in Osteoarthritis (OCTA) Research: Progress Update and Call for Membership Organizer: Tom Turmezei
11.00 – 12.15	**Session 3: "Imaging of pain mechanisms in OA"**
11.00 – 11.30	Keynote speaker: Dorothee Auer Moderators: David Hunter & Patrick Omoumi
11.30 – 12.15	Three abstracts Liu: The relationship between periarticular muscle properties and knee pain in non-overweight post-menopausal women Guermazi: The role of reader training, calibration and consensus sessions in the centralized, single-read, radiographic assessment paradigm used to determine patient eligibility in three phase III trials of Tanezumab for OA Carrino: Joint safety subgroup analysis in three phase III studies of Tanezumab
12.15 – 12.30	Poster pitches
12.30 – 13.00	Lunch & social gathering on Gather.Town
13.00 – 14.00	Poster session (odd poster numbers) @Gather.Town
14.00 – 15.00	**Session 4: "KNee OsteoArthritis Prediction (KNOAP2020) Challenge"**
14.00 – 14.30	Keynote speaker: Jukka Hirvasniemi Moderator: Stefan Klein
14.30 – 15.00	Presentation by challenge winners: Haresh Rajamohan: Deep learning for prediction of incident symptomatic knee OA
15.00 – 15.30	Break & social gathering on Gather.Town
15.30 – 17.00	**Session 5: "Imaging of bone-cartilage interactions in OA"**
15.30 – 16.00	Keynote speaker: Sharmila Majumdar Moderators: Saija Kontulainen & Gun-Il Im
16.00 – 17.00	Four abstracts Watkins: Areas of altered [18F]NAF PET uptake in response to exercise show OA progression on MRI over 2 years Liu: Automatic detection of BMLs from knee MRI data from the OAI study Eijkenboom: 3D patellar shape is associated with radiographic and clinical signs of patellofemoral osteoarthritis Bayramoglu: Analysis of patellar bone texture for automatic detection of patellofemoral osteoarthritis
17.00 – 17.30	Break & social gathering on Gather.Town
17.30 – 19.00	**Session 6: "Different perspectives on OA imaging"** 3 Keynote lectures: - *Patient perspective*: Elizabeth Cottrell & John Edwards - *Clinical care perspective:* Carl-Johan Tiderius - *Clinical trial perspective:* Nancy Lane Moderators: Margreet Kloppenburg & Thomas Link

Friday July 2, 2020

All times are Rotterdam time (CET)

9.30 – 10.45	**Session 7: "Big data OA imaging studies"**
9.30 – 10.00	Keynote speaker: Meike Vernooij Moderators: Sven Nebelung & Tom Turmezei
10.00 – 10.45	Three abstracts Wirth: MRI-based semiquantitative assessment (MOAKS) allows target selection of knees with accelerated quantitative cartilage thickness loss: Data from FNIH biomarker consortium Anwari: Knee pain is associated with perfusion kinetics at the infrapatellar fat pad among non-overweight postmenopausal women Rondas: The association between hip pain and radiographic hip osteoarthritis in the CHECK cohort
10.45 – 11.15	Break & social gathering on Gather.Town
11.15 – 12.45	**Session 8: "Image analysis/Artificial Intelligence in OA imaging"**
11.15 – 11.45	Keynote speaker: Simo Saarakkala Moderators: Erik Dam & Mikael Boesen
11.45 – 12.45	Four abstracts Felfeliyan: Domain adaptation on OAI dataset for unsupervised segmentation of bone and cartilage Rytky: Application of super-resolution deep learning for clinical cone-beam CT of ankle joint Nguyen: Interpretable deep learning framework for knee osteoarthritis structural prognosis prediction Caliva: Virtual bone aging: Can we predict 48 months of bone surface changes?
12.45 – 13.00	Poster pitches
13.00 – 13.30	Lunch & social gathering on Gather.Town
13.30 – 14.30	Poster session (even poster numbers) @Gather.Town
14.30 – 16.00	**Session 9: "Integrating biomechanics and imaging in OA"**
14.30 – 15.00	Keynote speaker: Janet Ronsky Moderator: James (JD) Johnston & Neil Segal
15.00 – 16.00	Four abstracts Jansen: Knee joint distraction results in MRI cartilage thickness increase up to ten years after treatment Bzowey: Focal changes in T2 relaxation time of loaded cadaver femoral cartilage Nebulung: Functional imaging of cartilage by serial T1ρ mapping - Different loading regimes in the assessment of tissue functionality Huppertz: Comprehensive serial T2 mapping to evaluate and monitor post-traumatic human cartilage degradation
16.00 – 16.30	Break & social gathering on Gather.Town

16.30 – 18.00	**Session 10: "The 'open data' paradigm"**
16.30 – 17.30	Panel presentations: OAI (Kent Kwoh), OA Trial Bank (Stefan Lohmander), Industry (Olga Kubassova), Platform developer (Serena Bonaretti), and Funding agency (Nelleke Richters; Dutch Arthritis Society, ReumaNederland) Moderators: Martin Englund & Frank Roemer
17.30 – 18.00	Panel discussion
18.00 – 18.30	**Adjourn and invitation to IWOAI 2022:** Nobutake Ozeki (Tokyo/Japan)
19.30 – 23.00	Gala dinner. Join us at https://ssrotterdam.nl/

THIS MEETING IS SUPPORTED BY A GRANT FROM PFIZER INC.

GOLD SUPPORT

BICL
BOSTON IMAGING CORE LAB

SILVER SUPPORT

BIOCLINICA®
biosplice
CALYX™
flexion
Transformative Medicine...
Where It Matters
ICM
image biopsy lab
KOLON TISSUEGENE
NOVARTIS

BRONZE SUPPORT

chondrometrics
DutchArthritisSociety
ReumaNederland
FUJIFILM
Value from Innovation
HAPLN SCIENCE
Aged peoples' Health & Happiness
IAG
IMAGE ANALYSIS GROUP
Organogenesis
Empowering Healing

ORAL PRESENTATIONS

KELLGREN AND LAWRENCE GRADE 2 AND 3 KNEES EXHIBIT A HETEROGENEOUS SPECTRUM OF CARTILAGE DAMAGE: THE MOST STUDY

*,**Roemer F.W., *Felson D.T., ***Stefanik J.J., *Rabasa G., *Wang N., *Crema M.D., *Neogi T., ****Nevitt M.C., *****Torner J., ******Lewis C.E., *Guermazi A.

*Boston University, Boston, MA, USA
**Friedrich-Alexander University Erlangen-Nürnberg (FAU) and Universitätsklinikum Erlangen, Erlangen, Germany
***Northeastern University, Boston, MA, USA
****University of California, San Francisco, CA, USA
*****University of Iowa, Iowa City, IA, USA
******University of Alabama at Birmingham, Birmingham, AL, USA

INTRODUCTION: Imaging plays an important role in defining structural osteoarthritis (OA) disease severity and potential suitability of patients to be recruited to disease-modifying OA drug (DMOAD) trials. Structural eligibility for clinical trials in OA usually is limited to KL2 and KL3 knees, knees that would ideally be similar in having damaged cartilage with sufficient remaining cartilage to be responsive to chondroprotective agents. KL2 knees are defined primarily by the presence of osteophytes, and may exhibit absence of cartilage damage and thus, are unlikely to benefit from cartilage-anabolic approaches. On the other hand, both KL2 and KL3 knees may have wide-spread cartilage loss to bone, making them potentially less responsive to a treatment focused on preserving cartilage. Understanding the proportion of knees and compartments without cartilage damage or with end-stage damage in a sample of KL2 and 3 knees will inform the decision-making regarding patient eligibility considerations.

OBJECTIVE: To describe frequencies of compartments without, with any, and with wide-spread full-thickness cartilage damage in OA knees with Kellgren-Lawrence (KL) grades 2 and 3.

METHODS: The Multicenter Osteoarthritis (MOST) study is a cohort study of individuals with or at risk for knee OA. All baseline MRIs with radiographic disease severity KL 2 and 3 were included. Knee MRIs were read for cartilage damage in 14 subregions. The frequencies of compartments without and with widespread full-thickness cartilage damage were described.

RESULTS: 445 knees had a radiographic disease severity of KL2 and 317 knees were KL3. Ninety-three (20.9%) KL2 knees had no damage in the medial (m) tibio-femoral joint (TFJ), 183 (41.1%) in the lateral (l) TFJ and 66 (14.8%) in the patello-femoral joint (Figure 1). (KL 3: 18/5.7%, 126/39.8% and 36/11.4%). 113 (24.8%) KL2 knees had wide-spread full-thickness damage in the medial TFJ, 48 (10.5%) in the lateral TFJ and 214 (47.0%) in the PFJ. (KL3: 235 (74.1%), 72 (22.7%) and 116 (36.6%)) (Figure 2).

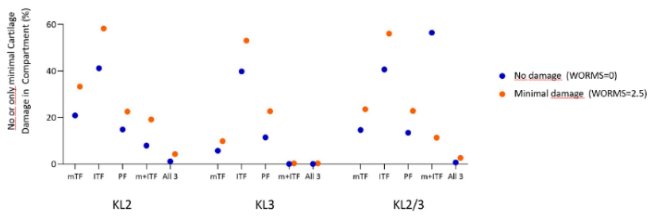


Fig. 1

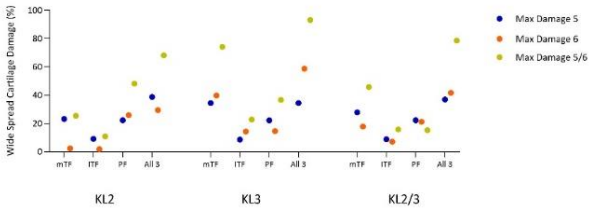


Fig. 2

CONCLUSION:
In summary, 20% of KL2 and 6% of KL3 knees do not exhibit any cartilage damage in the medial TFJ and one third of KL2 knees exhibits only minimal cartilage damage in the medial TFJ. Thus, these compartments without or only minimal damage are likely not assessable for cartilage-anabolic DMOAD effects. About 25% of KL2 and >70% of KL3 knees show wide-spread full-thickness cartilage damage in the medial TFJ and are likely not ideal candidates for anti-catabolic treatment approaches. For the lateral TFJ and the PFJ these numbers ranged from 10% (KL2, lateral TFJ) to 50% (KL2, PFJ). Given the heterogeneity of cartilage damage in KL2 and 3 knees, radiography as an instrument to define structural eligibility needs to be reconsidered depending on mode of action of a specific DMOAD compound.

SPONSOR: Supported by NIH grants from the National Institute of Aging to Drs. Lewis (U01-AG-18947), Torner (U01-AG-18832), Nevitt (U01-AG-19069), and Felson (U01-AG-18820).
DICLOSURE STATEMENT: AG is consultant for Merck Serono, AstraZeneca, Pfizer, Novartis, Regeneron, and TissuGene and is shareholder of BICL, LLC. FWR is shareholder of BICL, LLC. Dr. Roemer is consultant to Calibr.
CORRESPONDENCE ADDRESS: frank.roemer@uk-erlangen.de

ASSOCIATION OF SEMIQUANTITATIVE MRI-BASED MEASURES OF KNEE CARTILAGE WITH INCREASED CARTILAGE LOSS IN KNEES AT ELEVATED RISK OF DEVELOPING OA – DATA FROM THE OSTEOARTHRITIS INITIATIVE

*,**Wisser A., ***,****Roemer F.W., *,**Maschek S., *,**Eckstein F., ****Guermazi A., *,**Wirth W.

*Institute for Anatomy and Cell Biology, Paracelsus Medical University, Salzburg, Austria
**Chondrometrics GmbH, Freilassing, Germany
***Department of Radiology, University of Erlangen & BICL, Boston, MA
****Quantitative Imaging Center (QIC), Department of Radiology, Boston University, Boston, MA

INTRODUCTION: KLG0 knees with contralateral KLG2-4 (idiopathic OA, no trauma history) have been previously reported to exhibit greater cartilage thinning than KLG0 knees with contralateral KLG0, in particular in the lateral compartment. KLG0 knees with contralateral ROA could therefore be of interest for evaluating preventive DMOAD effects targeting early, pre-radiographic OA stages.

OBJECTIVE: The aim of the current study was to explore the association between baseline semi-quantitative (SQ) cartilage MOAKS scores and the three-year change in quantitative (Q) cartilage thickness in KLG0 knees with contralateral ROA. Further, we wished to explore, whether longitudinal change in cartilage thickness is associated with the severity of baseline MOAKS cartilage score parameters (i.e. area extent and full thickness involvement).

METHODS: The analysis was conducted in OAI participants with KLG0 in the target knee ('at risk knee') and ROA (KLG ≥ 2) in the contralateral knee. MOAKS assesses cartilage using two digits taking into account the area extent and the extent of full thickness loss in the same subregion (grades: 0.0, 1.0, 1.1, 2.0, 2.1, 2.2, 3.0, 3.1, 3.2 and 3.3). The maximum MOAKS area extent (MOAKSext, 1st component), the maximum MOAKS full thickness cartilage damage extent (MOAKSft, 2nd component) and the maximum MOAKS cartilage score (MOAKSmax) were determined in each 3 tibial and 1 central femoral MOAKS subregions. Medial / lateral compartment cartilage thickness (MFTC/LFTC) changes over 3 years were stratified by ipsi-compartmental baseline MOAKSmax, MOAKSext, and MOAKSft. Comparisons were performed using ANCOVA with adjustment for age, sex, and BMI and repeated with additional adjustment for the presence of effusion and adjacent BMLs, osteophytes, and meniscus damage/extrusion, to study the impact of other joint pathologies on the associations.

RESULTS: 150 knees of 150 participants were included (age 65±9 years (mean±SD), BMI 28±4 kg/m², 59% female). LFTC thickness loss (-0.03±0.15mm) tended to be greater than MFTC loss (-0.01±0.13mm) and 45% had MOAKS cartilage damage in the medial and 55% in the lateral compartment. Grades 1/2/3 of MOAKSext were present as follows: 13%/31%/2% medially, 17%/38%/0% laterally, and for grade 1/2 of MOAKSft: 7%/3% medially, 12%/7% laterally. MFTC cartilage thickness change did not differ for the various MOAKS strata compared to the without-lesion reference. LFTC cartilage thickness loss was higher in knees with presence of any lateral MOAKS lesion (adj. diff.: -0.06mm, 95% CI: [-0.11, -0.01]mm). This difference was mainly driven by the 46 knees with MOAKSmax 2.1/2.2. Knees with MOAKSext of 2 (adj. diff.: -0.08mm, [-0.13, -0.02]mm) and with MOAKSft damage (grade 1: -0.10mm, [-0.17, -0.02]mm; grade 2: -0.10mm, [-0.19, 0.00]mm) also showed greater LFTC loss than MOAKSext /MOAKSft 0 knees. Similar results were observed after adjustment for presence of other MOAKS features.

CONCLUSION: In this pre-ROA sample, presence of baseline MOAKS cartilage damage was associated with ipsi-compartmental cartilage loss over three years in the more commonly affected lateral compartment. Baseline MOAKS cartilage scores may therefore help selecting knees with expected subsequent cartilage thickness loss for inclusion in preventive DMOAD trials.

SPONSOR: Bundesministerium für Bildung und Forschung, Germany (01EC1408D)
DICLOSURE STATEMENT: F.E.: co-owner & CEO Chondrometrics GmbH; consulted for Merck, Bioclinica, Servier, Samumed, Roche, Kolon Tissuegene, Galapagos, Novartis & ICM; W.W.: CTO & co-owner Chondrometrics GmbH, consulted for Galapagos; A.W.: part-time employee Chondrometrics GmbH; S.M.: co-owner Chondrometrics GmbH. F.R. and A.G. are shareholder of BICL, LLC. A.G. is Consultant to Pfizer, Kolon TissueGene, Novartis, AstraZeneca, Merck Serono & Regeneron.
ACKNOWLEDGMENT: The OAI participants, OAI site investigators, coordinating center and funders
CORRESPONDENCE ADDRESS: anna.wisser@pmu.ac.at

EVALUATING THE TIBIAL SPINES AS IMAGING BIOMARKER FOR INCIDENT KNEE OA

*Runhaar J., *Damen J., **Oei E.H., *, ***Bierma-Zeinstra S.M.A.

*Erasmus MC University Medical Center Rotterdam, Department of General Practice, Rotterdam, the Netherlands
**Erasmus MC University Medical Center Rotterdam, Department of Radiology & Nuclear Medicine, Rotterdam, the Netherlands
***Erasmus MC University Medical Center Rotterdam, Department of Orthopedics and Sports Medicine, Rotterdam, the Netherlands

INTRODUCTION: Traditionally, radiographs are evaluated for osteophytes at the joint margins and joint space narrowing to evaluate OA disease status and as predictors for disease progression. For OA incidence, a KL grade of 1 is known to demonstrate strong predictive ability, however KL grading also known to be highly subjective and dependent on scoring/imaging methods. In the evaluation of radiographs, little attention has been paid to (spiking of) the tibial spines.

OBJECTIVE: This study explored the associations between different characteristics of the tibial spines on baseline radiographs and the incidence of knee OA after 30 months, among overweight, middle-aged women free of radiographic and clinical knee OA at baseline.

METHODS: We used data from the PROOF study. This RCT evaluated the effects of a lifestyle intervention and oral glucosamine on the incidence of knee OA. Given the absence of any treatment effects in the original RCT, data were handled as obtained from a cohort study. At baseline, 407 women, aged 50-60 years, with a BMI ≥ 27kg/m^2, free of knee symptoms and free of radiographic knee OA (KL grade <2) were enrolled. Baseline and 30 months follow-up measurements included a questionnaire to evaluate knee symptoms, and a standardized semi-flexed PA radiograph of both knees to evaluate KL grades, presence of spiking, and several measures of the medial and lateral tibial spine morphology (see Figure 1). A random set of 60 knee radiographs were additionally re-evaluated to calculate inter-observer reliability. Using Generalized Estimating Equations and a backward selection method (p > 0.1 for removal), Odds ratio (OR) and 95%CI were calculated for medial and lateral spiking, M°, L°, MT°, LT°, MH, LH, MH/W, and LH/W against the incidence of radiographic knee OA (KL grade ≥2) and against incident clinical knee OA (combined ACR criteria). For each outcome and for all factors with p ≤ 0.1, AUC-value and 95%CI against the outcome were calculated.

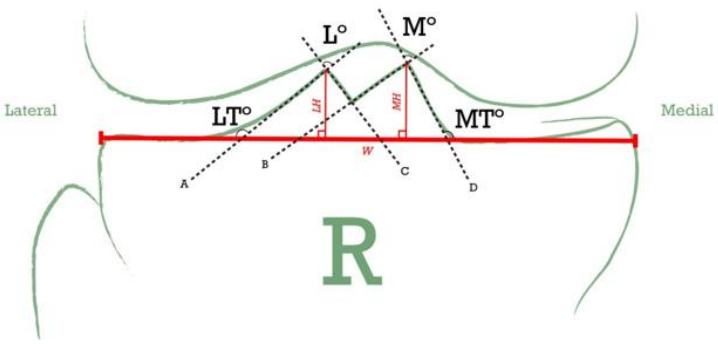


Figure 1. Measurement of tibial spine angulation and height in a right knee.
M° = Medial spine angle; L° = Lateral spine angle. The width of the tibial plateau (W) and the heights of the medial tubercle (MH) and lateral tubercle (LH) above the plateau were measured (in mm). The medial spine edge angle (MT°) is measured at the point of intersection between line W and line D. The lateral spine edge angle (LT°) is measured at the point of intersection between line W and line A.

RESULTS: 662 knees were available for the analyses. Mean baseline age was 55.7 ± 3.2 years and mean BMI was 32.0 ± 3.9 kg/m^2. ICC values for radiographic measures were high (≥0.7), apart from LT° (ICC 0.5). Only L° was significantly associated with incident radiographic knee OA (OR 0.98 [0.96-1.00]) and had an AUC of 0.63 (0.53-0.73). For incident clinical knee OA, medial spiking (OR 0.30 [0.10-0.91]), lateral spiking (OR 3.54 [1.41-8.90]), and MT° (OR 0.95 [0.91-1.00) were significantly associated, with an AUC of 0.71 (0.62-0.81).

CONCLUSION: Several measures of the tibial spines (i.e. lateral spine angle, presence of medial and lateral spiking, and medial spine edge angle) were significantly associated to incident radiographic and clinical knee OA, among a high-risk group of overweight, middle-aged women. These features could be considered in future prediction models, combined with other known predictors from history taking, physical examination, and/or imaging.

SPONSOR: N/A.
DICLOSURE STATEMENT: All authors have no disclosures.
ACKNOWLEDGMENT: Jeroen de Goffau (MD) is acknowledged for obtaining the tibal spine morphology measures.
CORRESPONDENCE ADDRESS: j.runhaar@erasmusmc.nl

DEEP LEARNING TO PREDICT EARLY RADIOGRAPHIC KNEE OA PROGRESSION DIRECTLY FROM MRI

*Panfilov E., **Nieminen M.T., **Saarakkala S., ***Tiulpin A.

*University of Oulu, Oulu, Finland
**University of Oulu & Oulu University Hospital, Oulu, Finland
***Aalto University, Espoo, Finland & Ailean Technologies Oy & University of Oulu, Oulu, Finland

INTRODUCTION: Existing hand-crafted knee MRI-based biomarkers, such as sub-regional cartilage or bone morphology, focus on major changes in these tissues. Design of the derived predictive models typically favors linear interactions of the features. This improves interpretability of the models but may limit the performance in predicting early knee OA changes. Recent works studied the power of deep learning (DL) in long-term structural progression or TKA prediction from MRI. However, applicability of DL in prediction of early progression remains uncertain. Such methods offer a powerful framework to efficiently leverage the available cohorts of multi-sequence MRI for discovery of composite imaging-based biomarkers.

OBJECTIVE: To investigate whether DL-based end-to-end approach can be used to predict radiographic OA progression over 12 and 48 months solely from MRI data.

METHODS: The complete baseline of the OAI was used in this study. We excluded subjects with KLG=4 at the baseline and the ones with missing KLG at 12- or 48-month exams. Subsequently, we selected only the subjects for which both sagittal 3D DESS and coronal IW TSE MRI scans were available. The progression criteria were defined as an increase in KLG at 12- and 48-month exams with no recorded decrease during the interval. Non-progression was defined as a persistent lack of change in KLG within the same period. The subjects matching neither of these criteria were excluded. All exams from one imaging center were chosen as a testing subset (n=1049/889, 1940/1637 knees for 12/48 months), the rest were used for training (n=3122/2722, 5739/4988 knees for 12/48 months). The training subset was used in 5-fold subject-wise cross-validation to develop an ensemble of 5 models. The reference logistic regression (LR) model performed classification into progressors/non-progressors from age, sex, BMI, and KLG at baseline. The imaging-based models were built independently for DESS and TSE using DL. The model architecture comprised an ImageNet-pretrained ResNet-50 encoder applied to sagittal (DESS) or coronal (TSE) slices, 2-layer LSTM for feature aggregation over the slices, and a fully connected layer to perform the final classification. The models were trained for 40 epochs to minimize focal loss, with the best snapshots identified by the highest balanced accuracy on validation. The reference and the DL models were compared in ROC AUC and average precision (AP) on the testing subset. Standard errors of the mean were estimated via bootstrapping (N=1000).

RESULTS: The LR showed ROC AUC and AP of 0.69±0.03 and 0.08±0.01 for 12 months, and 0.71±0.02 and 0.17±0.01 for 48 months intervals, respectively. For DESS, ROC AUC and AP were 0.69±0.03 and 0.08±0.01 for 12 months, and 0.71±0.02 and 0.23±0.02 for 48 months, respectively. For TSE, ROC AUC and AP were 0.65±0.03 and 0.07±0.01 for 12 months, 0.70±0.02 and 0.21±0.02 for 48 months.

CONCLUSION: Our results indicate that DL has a potential to identify the knees under a risk of radiographic OA progression from MRI data within 48 months. Here, the DL-based models may yield larger positive predictive values than the reference LR model. However, for progression within 12 months the observed performances were similar to the reference model ones. Interestingly, the DESS model showed higher predictive power than the TSE one in the 48 months case. Future work could investigate the differences of single-protocol models and the benefits of multi-sequence fusion in achieving even higher performance. Furthermore, analyzing the visual cues used by DL models may expose new informative MRI-based biomarkers associated with OA progression.

SPONSOR: Strategic funding of University of Oulu, Infotech Oulu.
DISCLOSURE STATEMENT: A.T. is a co-founder of Ailean Technologies Oy, a medical technology company providing consulting services.
ACKNOWLEDGMENT:
CORRESPONDENCE ADDRESS: egor.panfilov@oulu.fi

COMPOSITIONAL MRI OF ARTICULAR CARTILAGE OF THE WRIST AT 3T AND 7T MRI: A COMPARATIVE STUDY OF T2 and T2* RELAXOMETRY

*Heiss R.,**Weber M.A.,*,***Nagel A.M.,****Arkudas A., ****Horch R.E., *Hinsen M., *****Guermazi A., *Uder M., *,*****Roemer F.W.

*Department of Radiology, University Hospital Erlangen, Erlangen, Germany
**Institute of Diagnostic and Interventional Radiology, University Medical Center Rostock, Rostock, Germany
***Medical Physics in Radiology, German Cancer Research Center (DKFZ), Heidelberg, Germany
****Department of Plastic and Hand Surgery, University Hospital Erlangen, Erlangen, Germany
*****Department of Radiology, Boston University School of Medicine, Boston, MA, USA

INTRODUCTION: Compositional MRI techniques are being developed aiming to detect ultra-structural cartilage changes prior to macroscopic lesions. T2 is the most widely used compositional technique reflecting water and collagen content and collagen organization. T2 has already been applied in the ankle joint, knee, and hip and may be of interest in the evaluation of thin cartilage layers such of the wrist articulations. Similar to standard T2, T2* is assumed to similarly reflect cartilage composition based on interactions between water molecules and the collagen fibril network. Technical feasibility of biomechanical techniques to characterize pathology of the triangular fibrocartilage complex (TFCC), an important anatomical structure stabilizing the ulnar wrist, has been shown for 3T MRI. Ultra-high field MRI offers potential advantages compared to standard clinical system at 1.5 or 3T due to increased signal to noise ratio and potential higher spatial resolution. A direct comparison of 3T and 7T MRI for evaluation of multiple wrist cartilage surfaces and the TFCC has not been performed.

OBJECTIVE: To directly compare T2 and T2* relaxometry for evaluation of multiple wrist cartilage surfaces and the TFCC at 3T and 7T MRI.

METHODS: 50 subjects were enrolled, including a balanced patient (n=25) and a healthy volunteer group (n=25). Patients suffered from chronic wrist pain with suspected or confirmed osteoarthritis (OA) and/or ligament instability. The exclusion criteria were a history of trauma within the last 6 months, a history of inflammatory arthritis and inability to undergo 3T or 7T MRI. All participants of the study had both 3T and 7T MRI examinations. T2 MESE and T2* MESE sequences were acquired. Wrist cartilage surfaces were segmented in 6 ROIs, the TFCC was subdivided in 3 ROIs (Figure 1). Mean and standard deviations (SD) were determined. Spearman rho correlations were performed for T2 vs. T2* for both field strengths and for 3T vs 7T for T2 and T2*. Chi square statistics were used to evaluate differences between groups.

RESULTS: The patient group (n = 25) included 11 females and 14 males, with a mean age of 39.0 (range: 18–76). The volunteer group (n = 25) included 16 females and 9 males, with a mean age of 25.1 (age range: 16–34). Overall T2 values were higher for 7T compared to 3T for hyaline cartilage ROIs (range 34.7-46.5 ms for 3T and 40.9-53.7 ms for 7T, $p<0.01$ for 3/6 ROIs) and TFCC (range 24.7-33.3 ms for 3T and 31.2-41.2 ms for 7T, $p<0.05$ for 2/3 ROIs). SDs were higher for 7T (range 11.2-43.4 ms for 7T vs. 6.6-13.8 ms for 3T, $p<0.05$ for 9/9 ROIs). Correlations for T2 and T2* were poor or non-existent. (Table 1).

Table 1. Spearman's rho for T2 vs. T2* mean values for different ROIs (patients and controls)

	ROI1	ROI2	ROI3	ROI4	ROI5	ROI6	ROI7	ROI8	ROI9
Spearman's rho for mean values 3T patients	0.23	-0.42	0.07	0.22	0.23	0.14	-0.03	0.22	-0.38
Spearman's rho for mean values 7T patients	0.23	-0.32	0.39	0.16	-0.07	-0.30	-0.21	-0.25	-0.16
Spearman's rho for mean values 3T controls	0.26	0.17	0.43	0.52	0.50	-0.06	0.20	-0.03	0.56
Spearman's rho for mean values 7T controls	-0.06	0.07	0.52	0.44	-0.06	-0.02	-0.24	0.02	0.23

ROIs 1-5 and 9: hyaline cartilage; ROIs 6-8 TFCC

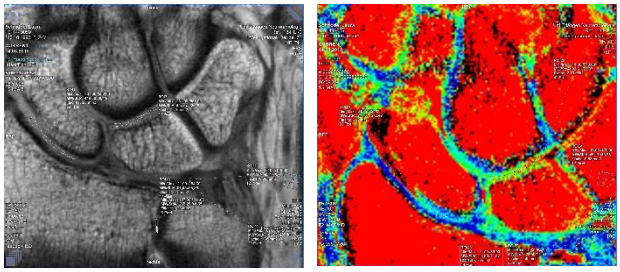


Fig 1. Example of T2 segmentation and color-coded T2 map

CONCLUSION: Unexpectedly T2 and T2* values in wrist cartilage plates and the TFCC are consistently higher for 7T compared to 3T. SDs were also higher for 7T reflecting increased noise, either due to technical issues (coil) or artifacts (e.g. motion). Correlations between T2 and T2* were low or absent reflecting that these compositional techniques do not reflect identical biochemical information.

SPONSOR: Supported by a research grant of the German Roentgen Society (Deutsche Röntgengesellschaft –DRG).
DICLOSURE STATEMENT: AG is Consultant to MerckSerono, AstraZeneca, Pfizer, Novartis, Regeneron, and TissuGene and is Shareholder of Boston Imaging Core Lab (BICL), LLC. FWR is Shareholder of BICL, LLC and Consultant to Calibr. None of the other co-authors declared potential competing interests.
CORRESPONDENCE ADDRESS: rafael.heiss@uk-erlangen.de

PILOT IMAGING *EX VIVO* HUMAN MENISCUS WITH PHASE-CONTRAST ENHANCED SYNCHROTRON MICRO-TOMOGRAPHY

*,**Einarsson E., ***Pierantoni M., ****Novak V., *,*****Svensson J., ***Isaksson H., **Englund M.

*Department of Translational Medicine, Lund University, Malmö, Sweden
**Department of Clinical Sciences Lund, Lund University, Lund, Sweden
***Department of Biomedical Engineering, Lund University, Lund, Sweden
****Swiss Light Source, Paul Scherrer Institute, Villigen, Switzerland
*****Medical Imaging and Physiology, Skåne University Hospital, Lund, Sweden

INTRODUCTION: An advantage with the use of synchrotron light is the high radiation and coherence which enable fast imaging with high resolution. Additionally, the use of phase contrast has the benefit of achieving high contrast in soft tissues.

OBJECTIVE: We aimed to investigate the feasibility of performing synchrotron radiation-based phase-contrast micro-tomography of *ex vivo* meniscus. Thus, in this pilot study, for the first time we obtained high-resolution images with the purpose to visualize and describe collagen structures of the human meniscus.

METHODS: We used medial meniscus samples from the MENIX biobank. The samples (n=4) were harvested from deceased donors, three women and one man, 18, 43 and 74 and 77 years old, respectively, without known knee OA. A 4mm radial slice was cut from the central body and stored frozen in phosphate buffered saline until imaging. Imaging of the thawed samples were conducted at the X02DA TOMCAT beamline at the Swiss Light Source, Switzerland, using a pco.edge 5.5 camera, with FOV 4.2 x 3.5mm, exposure time of 33ms, energy 15keV, 2001 projections over 180°, and propagation distance 150mm. The reconstructed images had a voxel side length of 1.63µm. Two volumes of $0.5x0.5x0.5mm^3$ from each sample, approximately at the center of the inner and middle zone of the meniscus respectively, were chosen for the image analysis (Fig. 1a). We calculated orientation maps using the structure tensor. Also, three texture features; contrast, correlation and energy, were calculated using gray level co-occurrence matrices (GLCMs) that were summed over four directions (0, 45, 90 and 135°) and the slices of the volume.

RESULTS: Collagen fibers could be clearly seen in the images (Fig. 1b). Calculated orientation maps further visualized the fiber direction (Fig. 1c). Main direction of the fibers was, as expected, circumferential. Images of one of the samples showed clear degeneration in the shape of fraying at the tissue border. A clear difference could be seen in the tissue structure of the inner zone of this sample compared to the others (Fig. 2). This sample was excluded from the subsequent analysis. The calculation of texture features resulted in high correlation values (mean values 0.82 and 0.83 in the inner and middle zone respectively) and low energy (mean values at the order of 10^{-3}) values, reflecting that the healthy meniscus is a highly structured and heterogeneous tissue. A somewhat lower contrast value in the middle zone (mean value of 60 compared to 77 in the inner zone) could reflect that the collagen fibers are more strictly ordered there compared to the inner zone.

CONCLUSION: Synchrotron radiation-based phase-contrast imaging is a promising tool to investigate *ex vivo* meniscus without the use of tissue fixation or embedding. It may be particularly useful to gain new insights in collagen fiber degradation and disorientation in OA, and to capture the tissues' response to different mechanical loads.

SPONSOR: Vinnova, Sweden.
DICLOSURE STATEMENT: None declared.
ACKNOWLEDGMENT: The Paul Scherrer Institut, Villigen, Switzerland, for provision of synchrotron radiation beamtime at the TOMCAT beamline.
CORRESPONDENCE ADDRESS: martin.englund@med.lu.se

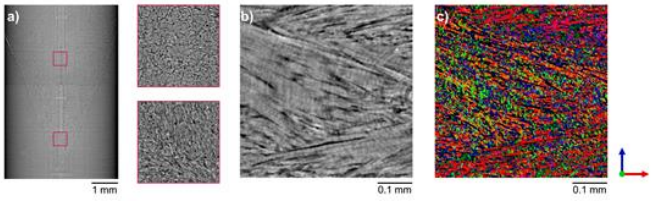


Fig. 2 Example images of the placement of two regions of interest at the center of the meniscus middle and inner zone, respectively (a), a slice of the chosen volume (middle zone) viewed from the tip of the meniscus (b) and an orientation map were fibers directed horizontally are colored red, vertically blue and through the image plane green (c).

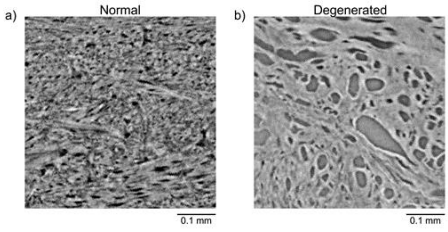


Fig. 1 Example images from the inner zone of a normal meniscus (a) and a visually degenerated meniscus (b).

3D GRADING OF CALCIFICATIONS FROM EX VIVO HUMAN MENISCUS POSTERIOR HORN

*Kestilä I., *Karjalainen V-P., *Finnilä M.A.J., **Folkesson E., **Turkiewicz A., **Önnerfjord P., **Hughes V., **Tjörnstrand J., **Englund M., *Saarakkala S.

*University of Oulu, Oulu, Finland
**Lund University, Lund, Sweden

INTRODUCTION: Meniscal calcifications are known to be associated with knee OA pathogenesis and could be a potential target for disease-modifying agents. Their distinct role in the disease process, however, is not well understood. In OA research, thin histological sections are typically used in the visualization of meniscal calcifications, although conventional histology is limited to 2D characterization. In this study, we propose a µCT-based 3D analysis of meniscal calcifications, including a new 3D grading system.

OBJECTIVE: 1) To introduce a novel µCT-based 3D grading of meniscal calcifications ex vivo, and 2) to describe the relationship between histopathological degeneration and the new calcification grade.

METHODS: Human medial and lateral meniscus posterior horns (n=40) were obtained from 10 TKA patients with medial compartment OA and 10 cadaveric donors representing reference population with no OA diagnosis. The samples were fixed in formalin; one subsection was treated with ascending ethanol concentrations and hexamethyldisilazane (HMDS), and imaged with µCT (SkyScan 1272, Bruker microCT; 50 kV, 200 µA, 2.0 µm pixel size, 2100 ms exposure time, Al 0.25mm filter). Adjacent to the µCT piece, thin tissue sections were obtained for histological stainings and Pauli's histopathological scoring. Calcifications were carefully inspected from the reconstructed 3D µCT images, and a 3D grading system was developed by modifying a histology-based system by Sun et al. (2010). Samples were graded from 0 to 5 according to newly developed grading criteria detailed in Figure 1.

RESULTS: 3D µCT images depict calcifications in all the samples. However, especially the minor calcifications are not visible in the histological sections (Fig. 1A). The calcification grades appear to be the highest in the medial menisci of medial compartment knee OA patients (Fig. 1B).

CONCLUSION: 3D µCT grading of meniscal calcifications is feasible. We found that µCT can depict more calcifications in 3D compared to conventional 2D histology: in the Alizarin red stained histological sections, the calcifications are visible only in the most severe cases. We believe that especially the smallest calcifications might get torn off from the tissue during the cutting process. The calcifications seem to increase with increasing severity of OA. Interestingly, it also seems that there are two types of degeneration processes in the meniscus: 1) without (or with a small number of) calcifications, and 2) with calcifications.

SPONSOR: Finnish Foundation for Technology Promotion, Grant No. 6562
DICLOSURE STATEMENT: none
ACKNOWLEDGMENT: We would like to thank the MENIX clinical staff at Trelleborg Hospital, the Tissue Donor Bank at Skåne University Hospital, and the Department of Forensic Medicine in Lund for their collaboration that enabled sample collection. We would also like to acknowledge Ms. Tarja Huhta, Laboratory Technician, for preparatory work with the histological samples.
CORRESPONDENCE ADDRESS: iida.kestila@oulu.fi

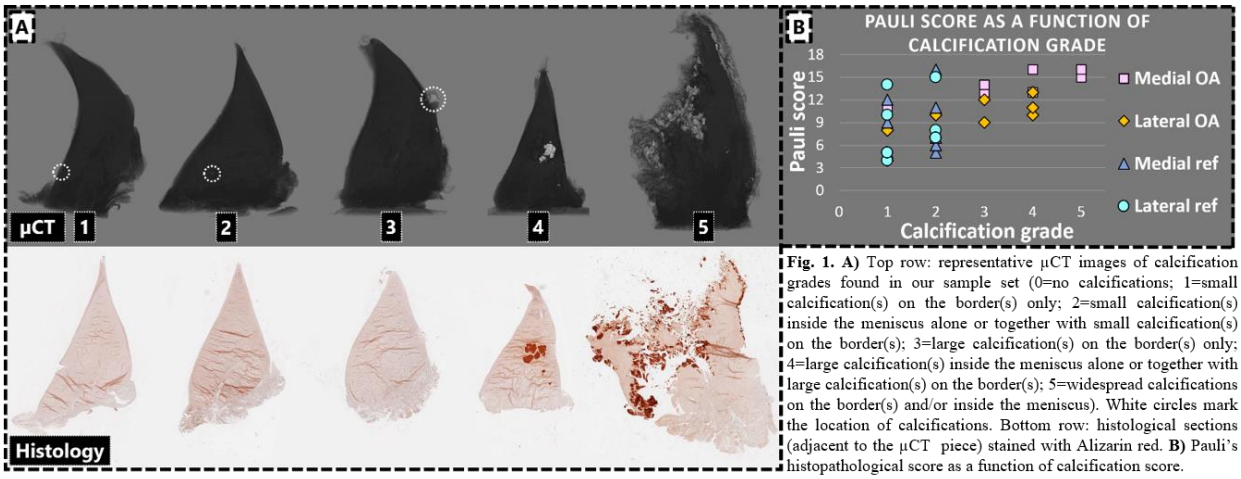


Fig. 1. A) Top row: representative µCT images of calcification grades found in our sample set (0=no calcifications; 1=small calcification(s) on the border(s) only; 2=small calcification(s) inside the meniscus alone or together with small calcification(s) on the border(s); 3=large calcification(s) on the border(s) only; 4=large calcification(s) inside the meniscus alone or together with large calcification(s) on the border(s); 5=widespread calcifications on the border(s) and/or inside the meniscus). White circles mark the location of calcifications. Bottom row: histological sections (adjacent to the µCT piece) stained with Alizarin red. **B)** Pauli's histopathological score as a function of calcification score.

EVALUATION OF AN AI SYSTEM FOR KNEE OSTEOARTHRITIS

*Egnell L., *Nexmann A., *Lems M., *Djernæs K., *Lisouski P., *Axelsen M.C., *Lundemann M.J.

*Radiobotics Aps, Copenhagen, Denmark

INTRODUCTION: As life expectancy continues to increase, the prevalence of OA is expected to increase even further, resulting in a heavier workload on general practitioners and radiology departments. In addition, the radiological interpretation and scoring is prone to unintended subjective variation often exacerbated by overly busy working environments among the reporting personnel. Here we evaluate a CE-marked medical device software, based on artificial intelligence (AI), that is capable of reading and reporting on knee OA, designed to support diagnostic decisions and aid the clinical workflow.

OBJECTIVE: The objective was to evaluate the standalone performance of a new device for radiographic knee OA analysis and reporting using two large, clinically representative datasets from the US.

METHODS: The device (RBknee™, Copenhagen Denmark) was separately tested on a subset of the OAI (422 subjects/4,611 imaged knees) and on the entire MOST dataset (3,024 subjects/20,113 imaged knees). All subjects (45–79 years old) were previously unseen and images were bilateral PA knee radiographs from baseline and follow-ups. Consensus reading by an expert panel on KLG, grading of JSN, osteophytes and subchondral sclerosis according to the OARSI atlas and mJSW were used as reference standards. The ability to identify the presence or absence of radiographic knee OA and related findings as well as to measure the minimum JSW (mJSW) was assessed using the following statistics; KLG and OARSI graded findings: sensitivity, specificity, and area under the receiver operating characteristic curve (AUC); mJSW: mean average error (MAE), intercept and slope by orthogonal linear regression.

RESULTS: The output was in strong agreement with the expert panel readings demonstrating an overall high performance (Table 1). In particular, the device detected radiographic OA (KLG≥2) with high sensitivity (0.88/0.97; OAI/MOST), high specificity (0.87/0.75), and high AUC (0.95/0.97). For the mJSW, for the medial and lateral compartment; MAE was 0.28 (95% CI: 0.27–0.30) and 0.39 (0.37–0.42), and the intercept/slope were -0.07/1.0 and -0.1/1.0 (OAI only).

CONCLUSION: This robust evaluation and the regulatory approval of the device paves the way for clinical implementations. While safety and effectiveness have been validated, a further prospective study could explore the implications of implementing such a tool in clinical practice.

SPONSOR: Radiobotics Aps, who is also the developer of the CE-marked device.
DISCLOSURE STATEMENT: M.C.A. and P.L. are shareholders of Radiobotics Aps.
ACKNOWLEDGMENT: We thank the NIH, OAI and MOST and all the PI's, contributors, participants and sponsors who contributed to these excellent datasets.
CORRESPONDENCE ADDRESS: liv@radiobotics.com

Table 1. Performance evaluation on subjects from the OAI and MOST datasets.

	OAI				MOST			
	Samples *n* (knees/subjects)	AUC	Sensitivity (95% CI)	Specificity (95% CI)	Samples *n* (knees/subjects)	AUC	Sensitivity (95% CI)	Specificity (95% CI)
KLG[†]	4279/421	0.95	0.88 (0.85–0.91)	0.87 (0.84–0.89)	18,448/2,801	0.97	0.97 (0.96–0.99)	0.75 (0.72–0.79)
JSN[‡]	4279/421	0.92	0.87 (0.84–0.90)	0.82 (0.79–0.86)	19,401/2,816	0.91	0.87 (0.84–0.90)	0.79 (0.75–0.82)
Osteophytosis[‡]	2427/289	0.92	0.89 (0.87–0.91)	0.78 (0.72–0.84)	18,723/2,801	0.95	0.93 (0.91–0.95)	0.79 (0.75–0.83)
Sclerosis[‡]	2329/272	0.91	0.84 (0.80–0.87)	0.87 (0.84–0.90)	n/a	–	–	–

[†] KLG ≥ 2, [‡] OARSI-grade ≥1. Abbreviations: Area under the receiver operating characteristic curve (AUC), Confidence Interval (CI).

THE RELATIONSHIP BETWEEN PERIARTICULAR MUSCLE PROPERTIES AND KNEE PAIN IN NON-OVERWEIGHT POST-MENOPAUSAL WOMEN

*,**Liu S., **Yazdankhah N., *,**Anwari V., *Tam K., Ha E., *Whyte R., **Naraghi A., **Sussman M.S., **Mohankumar R., ***Johnston J.D., **Probyn L., **Wong E., *MacKay C., *,**Rozenberg D.,*, **Wong A.K.

*University of Toronto, Toronto, Canada, ** University Health Network, Toronto, ON, Canada,
***University of Saskatchewan, Saskatoon, SK, Canada

INTRODUCTION: While being overweight is a KOA risk factor, between 15-40% of KOA patients are not overweight. There is a need to explore other KOA risk factors. It is known that lower muscle mass in the lower limbs and higher fat infiltration in the mid-thighs and mid-calves may relate to knee OA outcomes such as knee pain and functional limitation. For postmenopausal women in particular, estrogen depletion-related muscle loss and increased fat-related inflammation may be potential culprits for these knee OA outcomes. Currently, it is unknown whether muscles closer to the site of disease at the knee joint (periarticular muscles) better relate to knee pain and functional limitation. These muscles are already visible on knee MR images used for KOA management, but their role in KOA is unclear. Considering that MRIs are expensive, opportunistically examining muscle from the same knee MR images obtained for standard of care can add value to current clinical practice with minimal added cost.

OBJECTIVE: To examine how periarticular muscle (distal end of thighs, proximal heads of calves) properties (muscle volume, fat (inter- and intramuscular) volume and percentage (%)) relate to knee pain and function in non-overweight postmenopausal women (non-OW PMW). Hypotheses: H1) Greater fat% across all MR slices associates with greater knee pain and worse functional outcomes, H2) Greater muscle and fat volume, and fat% in MR slices closer to the knee joint have stronger associations with knee pain and functional outcomes.

METHODS: This cross-sectional study has recruited 55 participants from the community to date. Inclusion: non-OW PMW (BMI<25.0 kg/m^2, 50-85 years old) with varying degrees of knee pain. Exclusion: those with rheumatoid arthritis, existent joint replacements, or MRI contraindications. Participants' knees were scanned using a 3T MR scanner (quadrature knee coil) to capture muscle and fat. Axial and sagittal fat-saturated IM spin echo MRIs (repetition/echo times: 3433/15ms at 180^o flip angle, number of excitations=3, train length=7) were obtained with contiguous 2.0mm slices and in-plane resolution of 0.5x0.5mm spanning the prescribed regions of interest. Muscle and fat were segmented using a fully-automated iterative threshold-seeking algorithm developed by our lab using Python. Participants completed three questionnaires to assess various pain types (Knee Injury and Osteoarthritis Outcome Score (KOOS) pain subscale, Intermittent and Constant Osteoarthritis Pain, pain DETECT), and a stair climb, walk, and chair stand test to reflect knee function. Multivariable linear regression models determined how knee pain scores and functional measures each relate to 1) muscle properties summated across all slices and 2) muscle properties in each slice.

RESULTS: Among the 55 participants recruited, 46 participants had available MR images for image analysis (mean age=61.8$\pm$8.5 years, mean BMI= 22.44$\pm$3.20 kg/m^2). A 10% greater total fat% across all MR slices analyzed for the calf was associated with greater KOOS knee pain (B=-14.99 (p=0.037)) and greater intermittent pain (B=20.20 (p=0.030)) but no relationship with these measures was seen for total fat% of the thighs. For the thighs, slice-level analyses determined that most significant associations were found between muscle properties located in slices closer to the knee joint and functional outcomes. For the calf, most significant associations were found between muscle properties located in slices closer to the mid-leg and KOOS knee pain.

CONCLUSION: From preliminary results, overall fat infiltration in the proximal heads of the calf may be a potential risk factor for knee pain in non-OW PMW. Periarticular muscle information in slices closer to the knee joint for the thigh and closer to the mid-leg for the calf may be of greater importance to knee symptoms. To address potential false positive errors due to multiple testing, future steps include performing internal validation by bootstrapping at the patient and slice-level.

SPONSOR: Canadian Institutes of Health Research, Grant No PJT-156274.
DICLOSURE STATEMENT: No disclosures.
CORRESPONDENCE ADDRESS: andy.wong@uhnresearch.ca

THE ROLE OF READER TRAINING, CALIBRATION AND CONSENSUS SESSIONS IN THE CENTRALIZED, SINGLE-READ, RADIOGRAPHIC ASSESSMENT PARADIGM USED TO DETERMINE PATIENT ELIGIBILITY IN THREE PHASE III TRIALS OF TANEZUMAB FOR OSTEOARTHRITIS

*Guermazi A., *,**Roemer F.W., *Kompel A.J., *Diaz L.E., *,***Crema M.D., ****Brown M.T., ****Hickman A., *****Pixton G., ******Viktrup L., ****Fountaine R., ****Burr A., *******Sherlock S.P., ****West C.R.

*Boston University School of Medicine, Boston, USA
**Friedrich Alexander University Erlangen-Nürnberg & Universitätsklinikum Erlangen, Erlangen, Germany
***Institute of Sports Imaging, French National Institute of Sports, Paris, France
****Pfizer Inc, Groton, USA, *****Pfizer Inc, Morrisville, USA, ******Eli Lilly and Company, Indianapolis, USA
*******Pfizer Inc, Cambridge, USA

INTRODUCTION: Tanezumab is a nerve growth factor antibody in development for the relief of signs and symptoms of moderate to severe OA in adult patients for whom use of other analgesics is ineffective or not appropriate. After an FDA clinical hold, the tanezumab clinical trial program included rigorous x-ray-based eligibility assessment to exclude patients with, or at risk of, rapidly progressive OA (RPOA).

OBJECTIVE: To report the use of a single-read radiograph paradigm to support patient eligibility screening in three large, international, randomized, double-blind, Phase III studies of subcutaneous tanezumab (NCT02697773, NCT02709486 and NCT02528188).

METHODS: At screening, patients had bilateral shoulder, hip and knee radiographs taken by trained imaging technologists. These were read by one of five musculoskeletal radiology experts with no over-read. Factors for eligibility included confirmation of hip or knee OA, severity assessment by Kellgren Lawrence (KL) grading, identification of confounding joint conditions, the presence of RPOA, or potential risk factors for RPOA (such as osteonecrosis, subchondral insufficiency fracture, or atrophic OA). Central readers were trained using a program-specific imaging atlas and underwent a series of mock-reads to demonstrate image interpretation accuracy. Throughout the studies, inter-reader consistency was tracked with test cases blindly inserted into the reader queue. Readers attended quarterly calibration meetings where test discrepancies were discussed to achieve consensus and to align with the imaging atlas.

RESULTS: In the three studies, conducted at >480 international sites, 23,079 patients were screened and over 13,000 had radiographs assessed for eligibility (~84,000 radiographs). Across six sets of quarterly central reader testing, ≥4/5 trained radiologists agreed on the eligibility status of 73 to 90% of test cases (Table). Pair-wise reader agreement on overall radiographic eligibility ranged from 72 to 87%, with overall kappa for KL grading across the five readers ranging from 0.68 to 0.75 (Table 1).

Table 1. Inter-reader agreement on eligibility and Kellgren-Lawrence grading through 6 quarters of screening

	Eligibility						KL grading	
Testing quarter	**% agreement**	**Overall kappa**	**5/5 readers agree**	**4/5 readers agree**	**% ≥4/5 readers agree**	**3/5 readers agree**	**% agreement**	**Overall kappa**
1	72	0.43	12/30 test cases	12/30 test cases	80%	6/30 test cases	Not determined	0.70
2	77	0.52	15/30 test cases	10/30 test cases	83%	5/30 test cases	84	0.68
3	72	0.41	13/30 test cases	9/30 test cases	73%	8/30 test cases	77	0.68
4	77	0.51	16/30 test cases	7/30 test cases	77%	7/30 test cases	82	0.73
5	87	0.71	22/30 test cases	5/30 test cases	90%	3/30 test cases	77	0.75
6	77	0.45	18/35 test cases	10/35 test cases	80%	7/35 test cases	83	0.75

KL, Kellgren-Lawrence; Kappa, Fleiss' kappa.

CONCLUSION: A high rate of concordance can be achieved among central readers in a single-read radiograph paradigm when thorough training and ongoing assessment is employed. This method was used to support patient eligibility screening in three, large, multicenter, multinational clinical studies of subcutaneous tanezumab.

SPONSOR: These studies were sponsored by Pfizer and Eli Lilly and Company.
DICLOSURE STATEMENT: AG, FWR and MDC are shareholders at Boston Imaging Core Lab, LLC. AG has also been a consultant for AstraZeneca, MerckSerono, Pfizer, Organogenesis, Novartis, and TissueGene. FWR has also been a consultant for the California Institute for Biomedical Research. AJK and LED have no disclosures. MTB, AH, GCP, RJF, AB, SPS and CRW are employees of Pfizer and hold stock/stock options. LV is an employee of Eli Lilly and Company and holds stock.
ACKNOWLEDGMENT: The authors would like to acknowledge Bioclinica (Princeton, NJ, USA) for their central image processing expertise. Editorial support was provided by Jennifer Bodkin, PhD, CMPP, of Engage Scientific Solutions and was funded by Pfizer and Eli Lilly and Company.
CORRESPONDENCE ADDRESS: christine.west@pfizer.com

DEEP LEARNING FOR PREDICTION OF INCIDENT SYMPTOMATIC KNEE OSTEOARTHRITIS

*Rajamohan H.R., *Zhou Y., **Tan J., **Deniz C.M.

*Center for Data Science, New York University, New York, NY, USA
**Department of Radiology, New York University Langone Health, New York, NY, USA

INTRODUCTION**:** The prediction of symptomatic radiographic knee OA using clinical variables and knee radiographs is a challenging task. We developed a deep learning (DL)-based model to predict incident symptomatic knee OA (iSROA) within 78 months using baseline information from patients. Our method achieved the best performance at the Knee Osteoarthritis Prediction (KNOAP2020) Challenge.

OBJECTIVE**:** To develop a DL-based approach for predicting symptomatic radiographic knee OA.

METHODS**:** The training dataset was acquired from the OAI [1]. We included patients with the outcome variable, iSROA, available in the KNOAP2020 challenge. There were 2431 individuals and 3761 knee regions extracted from bilateral posteroanterior fixed-flexion knee radiographs using the method in [2]. The leaderboard OAI test dataset was provided with 108 cases. From the remaining 3653 cases in the OAI dataset, we chose individuals with age in the age range 47-65 and BMI in the range 23-50. This preselection is expected to achieve a minimal data distribution shift between the training, validation and the final KNOAP2020 test set. We performed a label-based (iSROA) stratified split into OAI training (80%, 1581 cases), OAI validation (20%, 403 cases) datasets.
The dataset had clinical variables such as Age, BMI, KL grade, injury history, tenderness etc. We selected Age BMI and KL grade for our analysis as the other clinical variables were missing for a subset of patients in the training and KNOAP2020 test set. We chose to use these 3 clinical features together with radiographic total knee replacement (TKR) predictions. We used pre-trained Resnet models [3] that predicted the likelihood of a patient undergoing a TKR within 9 years since we hypothesized that TKR would be correlated with the iSROA outcome. Ensembles of various numbers of TKR pretrained models were analyzed on the KNOAP2020 training set to identify the optimal ensemble DL-model based on AUC. The probability scores (post sigmoid) of the selected models were averaged (ensembled) and added as an additional feature to the patient knees (stacking). Using the selected features from clinical data (Age, KL, BMI) and TKR probability score from radiographs, we developed machine learning models - logistic regression (LR), two layers Multi-layer perceptron (MLP) and gradient boosted machine (GBM) with decision tree base learners to predict isROA using the OAI training set.

RESULTS: The ensemble of top 5 models from pretrained TKR models [3] achieved the best AUC of 0.84, when evaluated on the KNOAP2020 training set. On the OAI validation set, baseline LR model using only clinical variables achieved an AUC of 0.702±0.038. With the use of TKR probability scores, the LR model achieved an AUC of 0.717±0.038, which was comparable to the GBM (AUC = 0.716±0.040) and MLP (AUC = 0.715±0.039). The LR model achieved an AUC of 0.636 (95% CI: 0.571-0.699) on the final KNOAP2020 test data, and was the best performing model in the KNOAP2020 challenge

CONCLUSION**:** Predictive DL models pretrained on a related task can be used to improve the performance of models based on clinical variables for predicting iSROA within 78 months.

DISCLOSURE STATEMENT: None
ACKNOWLEDGMENT: The study was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) of the National Institute of Health (NIH) under award number R01 AR074453
CORRESPONDENCE ADDRESS: cem.deniz@nyulangone.org

REFERENCES:
[1] https://data-archive.nimh.nih.gov/oai/
[2] Zhang, Tan, Cho, Chang, & Deniz (2020). Attention-based CNN for KL grade classification: Data from the osteoarthritis initiative. In *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)* (pp. 731-735)
[3] Leung et al. (2020) Prediction of total knee replacement and diagnosis of osteoarthritis by using deep learning on knee radiographs: data from the osteoarthritis initiative. *Radiology* 296:584-593.

JOINT SAFETY SUBGROUP ANALYSES IN THREE PHASE 3 STUDIES OF TANEZUMAB

*Carrino J.A., **McAlindon T.E., ***Brown, M.T., ***Burr A., ***Fountaine R.J., ****Pixton G., *****Viktrup L., ***West C., ***Verburg, K.M.

*Department of Radiology, Hospital for Special Surgery, New York, NY, USA
**Department of Rheumatology, Tufts Medical Center, Boston, MA, USA
***Pfizer Inc., Groton, CT, USA
****Pfizer Inc., Morrisville, NC, USA
*****Eli Lilly and Co., Indianapolis, IN, USA

INTRODUCTION: Joint safety events are an uncommon but important part of the tanezumab (nerve growth factor antibody) safety profile in patients with moderate to severe OA.

OBJECTIVE: We conducted integrated subgroup analyses into the potential patient-level factors associated with joint safety events.

METHODS: All possible/potential joint safety events across 3 subcutaneous tanezumab trials in patients with moderate to severe hip or knee OA for whom other treatments were ineffective or not appropriate (placebo control: NCT02697773[1] & NCT02709486[2]; NSAID control: NCT02528188[3]) were adjudicated by a blinded external Adjudication Committee (AC). The adjudicated composite joint safety endpoint (CJSE) included primary osteonecrosis, rapidly progressive OA (RPOA) type 1 and 2, subchondral insufficiency fracture, and pathological fracture. Data were cross-tabulated to explore potential associations for CJSE, RPOA1 or RPOA2 with numerous baseline demographics, clinical characteristics, bone health, and post-baseline outcome variables (adverse events [AEs]; efficacy response; sensory examinations; nonsteroidal anti-inflammatory drug (NSAID) and non-NSAID concomitant medications). No formal statistical analyses were conducted.

RESULTS: 145 of the 4541 treated patients (3.2%) had an adjudicated CJSE; of these, 125 had RPOA1 or RPOA2. Exploratory subgroup analyses showed associations for CJSE, RPOA1 and RPOA2 with baseline Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC)* pain score ≥7/10 and emergent AEs of arthralgia and joint swelling; CJSE and RPOA1 with baseline WOMAC physical function ≥7/10 and maximum KL grade 3 (any joint); and RPOA2 with maximum KL grade 4 (any joint). There were no consistent associations with demographics, efficacy response, or concomitant medications.

CONCLUSION: Among patients with moderate to severe OA treated in 3 placebo- or NSAID-controlled trials of tanezumab, 3.2% had an adjudicated CJSE. Cross-treatment subgroup analyses showed associations for RPOA with baseline OA severity (symptomatic and radiographic) but not with demographics, efficacy response, or concomitant medications.

REFERENCES:
1. Schnitzer, T. et al. JAMA. 2019;322:37-48.
2. Berenbaum, F. et al. Ann Rheum Dis. 2020;79:800-10.
3. Hochberg, M. et al. Arthritis Rheumatol. 2021; Epub Feb 3.

SPONSOR: These studies were funded by Pfizer and Eli Lilly and Company.
DISCLOSURE STATEMENT: JAC has received consulting fees from Covera Health, Globus, and Pfizer, and is on the medical advisory board of Image Analysis Group, Image Biopsy Lab, and Carestream (non-paid positions). TEM has served on an advisory board for Pfizer, Samumed, Seikagaku, Flexion Therapeutics, Sanofi, Roche, Astellas, and Regeneron. MTB, AB, RJF, GP, CW, and KMV are employees of Pfizer Inc. and have stock/stock options. LV is an employee of Eli Lilly and Company and has stock.
ACKNOWLEDGMENT: Editorial support was provided by Jennifer Bodkin, PhD, CMPP, of Engage Scientific Solutions and was funded by Pfizer and Eli Lilly and Company.
CORRESPONDENCE ADDRESS: CarrinoJ@HSS.EDU

AREAS OF ALTERED [^{18}F]NAF PET UPTAKE IN RESPONSE TO EXERCISE SHOW OA PROGRESSION ON MRI OVER 2 YEARS

*Watkins L., **Mackay J., *Kogan F.

*Stanford University, Stanford, CA, USA
**University of East Anglia, Norwich, UK & University of Cambridge, Cambridge, UK

INTRODUCTION: Bone is a metabolically active tissue whose physiology can change quickly in response to mechanical loading. Abnormal bone physiologic response to a mechanical load thus may serve to detect improper whole-joint function which is relaying focally abnormal bone loads. [^{18}F]sodium fluoride ([^{18}F]NaF) is a well-established bone-seeking agent that detects regions of altered bone metabolism and acute loading of the knee with exercise altered [^{18}F]NaF uptake in healthy and OA subjects, suggesting this method may detect these areas of abnormal joint function. Further, hybrid [^{18}F]NaF PET/MR imaging can further combine this information with structural and microstructural soft-tissue information from MRI.

OBJECTIVE: We evaluated the relationship between the acute response of [^{18}F]NaF PET uptake to exercise and the progression of structural features of OA over 2 years as assessed by MRI.

METHODS: Three subjects with knee OA (54 ± 9 years; BMI 25.9 ± 0.7; 1 F) were scanned using a 3T whole-body hybrid PET-MRI system under an approved IRB protocol both at an initial timepoint and again after 2 years. At the initial timepoint, subjects were scanned twice on the same day: first at rest, then again after performing a 1-legged squat exercise activity to fatigue. Subjects performed the exercise on the affected knee, or on the knee with greater pain if bilateral knee OA was reported. Focal areas with abnormally high increases in post-exercise [^{18}F]NaF PET uptake (ROI_{focal}) were defined as 4 or more adjacent voxels with an absolute SUV increase after exercise greater than three. Subjects subsequently returned for an MRI-only examination 2 years after the initial assessment. At both timepoints, MRI Osteoarthritis Knee Score (MOAKS) assessment was performed to identify normal-appearing bone and cartilage (MOAKS 0) and regions with BML, OP, or cartilage lesions (MOAKS score > 0) across 8 bone/cartilage regions. Areas where the MOAKS score changed between timepoints were compared with ROI_{focal} from 2 years prior.

RESULTS: There were 14 ROIs with an abnormal focal response in [^{18}F]NaF uptake to exercise at the initial timepoint. Four of these areas appeared structurally normal on MRI, while 10 were associated with bone marrow lesions, osteophytes, and/or cartilage loss. MR images at the 2-year follow-up time showed that 14 regions had a change in MOAKS score: 11 were ROI_{focal}, while 3 regions were not noted as having any abnormal response to exercise at the initial timepoint. The rest of the joints were stable and did not have any changes in this time period. Ten of the ROI_{focal} already had a structural feature and significantly higher activity than the surrounding bone at the initial timepoint; 9 of them went on to have that feature change in size. There were 4 bone marrow lesions at the initial timepoint, 2 increased in size while the other 2 shrank in size but developed worsening cartilage loss and/or larger osteophytes. Other changes included progression or development of cartilage loss and/or osteophytes. Of the four 4 ROI_{focal} that appeared structurally normal at the initial timepoint, 2 went on to develop a structural finding (2 cartilage lesions and 1 bone marrow lesion) while the other two did not show any OA progression over 2 years.

CONCLUSION: We used an exercise protocol to acutely load the knee and examine the metabolic response of bone as a measure of joint function. After exercise, numerous focal regions of increased [^{18}F]NaF uptake were observed. While the majority represented structural OA features, there were some ROI_{focal} that showed no abnormalities in [^{18}F]NaF uptake compared to adjacent regions pre-exercise or structural bone abnormalities on MRI. Over three-quarters of ROI_{focal} showed structural changes 2 years later. These findings suggest that changes in [^{18}F]NaF PET uptake after exercise may help identify, at a single timepoint, regions where there are functional abnormalities in joint response to acute load that are at risk of OA progression.

SPONSOR: NIH grants R01AR074492, R21EB030180, R00 EB022634.
DISCLOSURE STATEMENT: The authors receive research support from GE Healthcare.
CORRESPONDENCE ADDRESS: lewatkin@stanford.edu

AUTOMATIC DETECTION OF BONE MARROW LESIONS FROM KNEE MRI DATA FROM THE OAI STUDY

*Liu S., *Sun X., **,***Roemer F.W., ****Ashbeck E.L., *Bedrick E.J., ****Li Z.M., ***Guermazi A., ****Kwoh C.K.

*Department of Epidemiology and Biostatistics, University of Arizona, AZ, USA
**Department of Radiology, University of Erlangen – Nuremberg, Erlangen, Germany
***Department of Radiology, Boston University School of Medicine, MA, USA
****The University of Arizona Arthritis Center, University of Arizona, AZ, USA

INTRODUCTION: Subchondral bone marrow lesions (BMLs) are associated with symptoms and structural progression of knee OA. BMLs are characterized by diffuse hyperintensity on fat suppressed T2-weighted images. BMLs are commonly assessed using semi-quantitative (SQ) scoring systems by expert readers or volume quantification. Automated detection of BMLs using machine learning approaches may help in screening potential participants in clinical trials enriched for fast structural progression.

OBJECTIVE: To automatically detect BMLs from MRI data using deep learning models.

METHODS: The dataset consists of 1,899 knee MRI exams from three sub-studies of the Osteoarthritis Initiative (OAI) with available baseline SQ readings: Foundation for NIH (FNIH) OA biomarker, Pivotal OAI MRI Analysis (POMA) incident radiographic knee OA, and POMA knee replacement. We trained a deep learning model (modified MRNet) using the sagittal intermediate-weighted (IW) fat-suppressed (FS) images. We dichotomized the MOAKS (MRI Osteoarthritis Knee Score) grades into presence or absence categories. The split was done by categorizing grades > 0 as presence and grades = 0 as absence. The whole data were randomly split into a training set (1524 exams) to train the model, a validation set (182 exams) to select the best model, and a test set (193 exams) to evaluate prediction performance. After the deep learning models were trained, we obtained probabilities of the existence of BMLs from IW images on each of 15 subregions in the femur and tibia. The logistic regression model was trained to combine the probabilities of different regions for IW images to output a probability of BMLs status for each subject overall. We utilized the area under the receiver operating characteristic curve (AUC) to assess the model's performance and 95% confidence intervals of AUC to measure the variability of AUC. AUC values range from 0 to 1, with 1 indicating a 100% correct detection.

RESULTS: The highest AUC value for IW images in the test set was 0.90 at two regions (tibia central medial and patella lateral). AUC values ranged from 0.51 (femur medial anterior) to 0.89 (femur lateral anterior) in the test set for the other regions. At the subject level, the detection performance of the combined model achieved an AUC value of 0.87 with a 95% confidence interval of [0.82, 0.93].

CONCLUSION: We have developed and validated a fully automated deep learning framework to detect BMLs from MRI data. After the model had been trained, we applied the trained model to MRI data without MOAKS readings to predict the probability of BMLs status automatically. The prediction performance of detecting BMLs varied among different regions. Our method has some limitations, e.g., the prediction accuracy at the subject level is not as good as the accuracy for the individual subregions. In future studies, we will utilize images from other MRI sequences and additional MOAKS readings to improve our model's performance. Automated classification of diffuse vs cystic BMLs and severity grading will be important in addition.

SPONSOR: None.
DICLOSURE STATEMENT: AG is consultant to Pfizer, Novartis, Regeneron, TissueGene, Merck Serono, and AstraZeneca. He is shareholder of BICL, LLC. FR is shareholder of BICL, LLC and consultant to Calibr –California Institute of Biomedical Research. KK is consultant to Regeneron, LG Chem, and Express Scripts. He is principal investigator for pharma sponsored clinical trials to Abbvie, UCB, Eicos, Cumberland, Mitsubishi, and GSK and DSMB to Kolon TissueGene and Avalor Therapeutics.
CORRESPONDENCE ADDRESS: xiaosun@arizona.edu

3D PATELLAR SHAPE IS ASSOCIATED WITH RADIOGRAPHIC AND CLINICAL SIGNS OF PATELLOFEMORAL OSTEOARTHRITIS

*Eijkenboom J.F.A., **Tümer N., *Schiphof D., ***Oei E.H., **Zadpoor A.A., *Bierma-Zeinstra S.M.A., *van Middelkoop M.

*Department of General Practice, Erasmus MC University Medical Center Rotterdam, The Netherlands
**Department of Biomechanical Engineering, Delft University of Technology, The Netherlands
***Department of Radiology & Nuclear Medicine, Erasmus MC University Medical Center Rotterdam, The Netherlands

INTRODUCTION: Several factors alter the mechanics of the patellofemoral (PF) joint, which are hypothesized to increase joint stresses, that can in turn lead to osteoarthritis (OA). Biomechanical factors as joint loading, gait and alignment have been broadly studied and associated with the presence of PFOA. However, little is known about 3D patellar bone shape.

OBJECTIVE: To examine the association between 3D patellar shape and 1) isolated radiographic PFOA, 2) morphological features of PFOA and 3) clinical symptoms of PFOA.

METHODS: Magnetic Resonance Imaging (MRI, 1.5T) data from 66 women with isolated radiographic PFOA and 66 age and BMI-matched controls were selected from a cohort study (Rotterdam Study (RS-III-1)). Mean age of the participants was 57 years (SD 3.9) with an average BMI of 27.5 kg/m^2. Patellae were manually segmented from MRI (fatsat) and subsequently used to build 3D statistical shape models of the patella. Structural abnormalities were semi-standardized scored on MRI using MOAKS and basic clinical symptoms of PFOA were documented. Regression analyses were applied to determine associations between the 3D independent shape modes and group status, clinical symptoms and structural abnormalities.

RESULTS: Four shape variants (mode 6, 13, 21 and 23) showed a statistical significant association with group status. Three of these variants showed an association with the presence of osteophytes and cartilage loss on the patella (Table 1). Two of these variants showed an association with clinical symptoms of PFOA.

CONCLUSION: Patellar shape is associated with the presence of radiographic PFOA in middle-aged women. Some shape variants were also associated with clinical symptoms. Interestingly, shape variant 6 was earlier shown to be associated with structural abnormalities associated with OA in a population aged under 40. This may suggest that patellar shape may be an early detectable risk factor for PFOA.

SPONSOR: Dutch Arthritis Association
DICLOSURE STATEMENT: No disclosures.
CORRESPONDENCE ADDRESS: m.vanmiddelkoop@eramsusmc.nl

TABLE 1. Associations between MOAKS features connected to PFOA and patellar shape modes (OR with 95% CI).

	Osteophytes patella	**Osteophytes anterior Femur**	**BMLs patella**	**BMLs anterior Femur**	**Cartilage loss patella**
Mode 6	**0,55 (0,37-0,84)**	0,68 (0,45-1,03)	**0,60 (0,40-0,90)**	0,83 (0,56-1,22)	**0,49 (0,32-0,75)**
Mode 13	**0,64 (0,43-0,97)**	0,74 (0,49-1,11)	**0,68 (0,46-0,99)**	**0,59 (0,39-0,90)**	**0,64 (0,44-0,95)**
Mode 21	0,71 (0,48-1,06)	**0,62 (0,41-0,94)**	0,86 (0,60-1,24)	0,76 (0,51-1,11)	**0,62 (0,42-0,91)**
Mode 23	**2,42 (1,53-3,84)**	1,38 (0,95-2,02)	1,07 (0,77-1,50)	1,13 (0,80-1,61)	1,32 (0,94-1,87)

Adjusted for age and BMI. Significant associations ($p<0.05$) in bolt.

ANALYSIS OF PATELLAR BONE TEXTURE FOR AUTOMATIC DETECTION OF PATELLOFEMORAL OSTEOARTHRITIS

*Bayramoglu N., *,**,***Nieminen M.T., *,**,***Saarakkala S.

*Research Unit of Medical Imaging, Physics and Technology, University of Oulu, Oulu, Finland
**Department of Diagnostic Radiology, Oulu University Hospital, Oulu, Finland
***Medical Research Center, University of Oulu and Oulu University Hospital, Oulu, Finland

INTRODUCTION**:** Patellofemoral osteoarthritis (PFOA) is highly prevalent and clinically important, yet it still under investigated. In the clinical setting, radiographic tibiofemoral (TF) joint is often evaluated with a posterior-anterior (PA) radiograph from which the PF joint cannot be evaluated at all. In addition, pure clinical assessment cannot be used either to diagnose patellofemoral OA. Since it has been reported that, at least in some phenotypes of knee OA, PF joint changes even precedes TF joint changes, there is a strong need to also consider PF joint when developing new imaging biomarkers for diagnosing and monitoring knee OA.

OBJECTIVE**:** To assess the ability of texture features for detecting radiographic PFOA from knee lateral view radiographs.

METHODS**:** We used lateral view knee radiographs from The Multicenter Osteoarthritis Study (MOST) public use datasets (*n* = 5507 knees). Patellar region-of-interest (ROI) was automatically detected using landmark detection tool (BoneFinder software), and subsequently, these anatomical landmarks were used to extract three different texture ROIs. Hand-crafted features, based on Local Binary Patterns (LBP), were then extracted to describe the patellar texture. First, a machine learning model (Gradient Boosting Machine) was trained to detect radiographic PFOA from the LBP features. Furthermore, we used end-to-end trained deep convolutional neural networks (CNNs) directly on the texture patches for detecting the PFOA. The proposed classification models were eventually compared with more conventional reference models that use clinical variables such as age, sex, body mass index (BMI), WOMAC score, and tibiofemoral KL grade from PA radiographs. Atlas-guided visual assessment of PFOA status by expert readers provided in the MOST public use datasets was used as a classification outcome for the models. Performance of prediction models was assessed using the area under the receiver operating characteristic curve (ROC AUC) and the area under the precision-recall (PR) curve - average precision (AP) - in the stratified 5-fold cross validation setting.

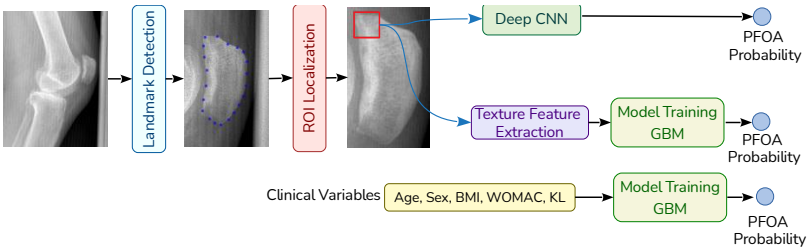


Figure1: Illustration of the workflow of our approach.

RESULTS**:** Of the 5507 knees, 953 (20%) had PFOA. AUC and AP for the strongest reference model including age, sex, BMI, WOMAC score, and tibiofemoral KLG to predict PFOA were 0.817 and 0.487, respectively. Textural ROI classification using CNN significantly improved the prediction performance (ROC AUC= 0.889, AP= 0.714).

CONCLUSION**:** We present the first study that evaluates patellar bone texture for detecting PFOA. Good classification performance values indicate that analyzed texture features contain useful information of patellar bone structure, which seems to change in PFOA. These texture features may be used in the future as novel imaging biomarkers in OA diagnostics.

SPONSOR: None
DICLOSURE STATEMENT: None
ACKNOWLEDGMENT: Claudia Linder and Aleksei Tiulpin for assistance in landmark localization.
CORRESPONDENCE ADDRESS: neslihan.bayramoglu@oulu.fi

MRI-BASED SEMIQUANTITIAVE CARTILAGE ASSESSMENT (MOAKS) ALLOWS TARGETED SELECTION OF KNEES WITH ACCELERATED QUANTITATIVE CARTILAGE THICKNESS LOSS: DATA FROM THE FNIH BIOMARKER CONSORTIUM

*Wirth W., *Maschek S., *Wisser A., **Guermazi A., ***Hunter D.J., *Eckstein F, ****Roemer F.W.

*Paracelsus Medical University, Salzburg, Austria & Chondrometrics GmbH, Freilassing, Germany
**Department of Radiology, Boston University & BICL, Boston, MA, USA
***Rheumatology Department, Royal North Shore Hospital and Institute of Bone and Joint Research, Kolling Institute, University of Sydney, Sydney, NSW Australia
****Department of Radiology, University of Erlangen & Boston University and BICL, Boston, MA, USA

INTRODUCTION: Structural eligibility in DMOAD trials usually relies on radiographic selection criteria (e.g., KLG or medial JSN). MRI-based MOAKS cartilage lesion scores may, however, allow to select knees with accelerated progression, which are considered ideal candidates for demonstrating DMOAD efficacy, more specifically than radiographic selection criteria.

OBJECTIVE: To study, whether and which baseline MRI-derived MOAKS cartilage scores are predictive for subsequent ipsi-compartmental accelerated cartilage thickness loss and how the area extent component of the MOAKS cartilage lesions and the extent of the full-thickness component of MOAKS cartilage assessment impact cartilage thickness loss. In addition, we assessed the sensitivity to change for MOAKS and medial JSN strata.

METHODS: All 600 participants from the OAI FNIH study had MOAKS assessments and quantitative cartilage thickness measurements (age: 62y, BMI: 31kg/m², 59% female). The maximum MOAKS cartilage score (MOAKSmax), the maximum MOAKS cartilage damage extent score (MOAKSext, 1st component of MOAKS cartilage scale), and the maximum MOAKS cartilage damage full thickness score (MOAKSft, 2nd MOAKS cartilage component) in the 3 tibial and the central femoral MOAKS subregion were determined for the medial and lateral compartment. Medial and lateral femorotibial compartment cartilage thickness (MFTC/LFTC) change over the subsequent two years were stratified by MOAKSmax, MOAKSext, and MOAKSft score in these knees. Between-group comparisons were performed using ANCOVA with adjustment for age, sex, and BMI and repeated with additional adjustment for the presence of effusion and adjacent BMLs, osteophytes, and meniscus damage/extrusion, to study the impact of other joint pathologies on the observed associations.

RESULTS: MFTC cartilage thickness loss was greater in knees with MOAKS cartilage lesions than in knees without (except from small superficial lesions and small strata), with and without additional adjustment for other MOAKS features. The greatest magnitude of change and the greatest sensitivity to change (SRM) were observed in knees with full thickness defects and in knees with large, widespread superficial defects (SRM range: -0.62 to -1.34). When stratifying by the extent of cartilage damage or extent of full thickness cartilage damage, SRMs increased with inceasing MOAKSext (grade 1/2/3: -0.45 / -0.61 / -0.73) as well as with increasing MOAKSft (grade 1/2/3: -0.76 / -0.89 / -0.95). In comparison, SRMs were -0.47 / -0.68 in knees with medial JSN 1 / 2. Interestingly, 67 of the 438 knees with medial JSN 1 / 2 had no medial cartilage damage (MOAKSmax of 0.0) and 98 of the 219 knees with medial JSN 2 had no full thickness lesions. In the lateral compartment, both the cartilage thickness loss and the MOAKS scores were notably lower than in the medial compartment (data not shown).

CONCLUSION: Presence of MOAKS cartilage damage was associated with the magnitude of cartilage loss. Presence of full thickness lesions, but also widespread superficial lesions may allow to specifically select knees likely to exhibit high sensitivity to change but also to exclude knees, in which a DMOAD is unlikely to be effective (e.g., wide-spread full thickness damage). Selection by MOAKS scores or simplified SQ measures such as ROAMES may therefore allow reduction of the sample size when compared to JSN grades, that were observed to not necessarily reflect cartilage damage.

DISCLOSURE STATEMENT: WW: Consulting for Galapagos; DJH: consulting for Pfizer, Lilly, TLCBio, Novartis, Tissuegene, Biobone; FE: consulting for Merck KGaA, Abbvie, Samumed, Kolon-Tissuegene, Servier, Galapagos, Roche, Novartis, ICM and HealthLink. AG: consulting for Pfizer, Kolon TissueGene, Novartis, AstraZeneca, Merck Serono and Regeneron.
ACKNOWLEDGMENT: Foundation for the NIH (FNIH), OAI participants, staff, and investigators.
CORRESPONDENCE ADDRESS: wolfgang.wirth@pmu.ac.at

KNEE PAIN IS ASSOCIATED WITH PERFUSION KINETICS AT THE INFRAPATELLAR FAT PAD AMONG NON-OVERWEIGHT POSTMENOPAUSAL WOMEN

*Anwari V., *Siwen Liu., *Yazdankhah N., *,**Ha E., *Whyte R., *Naraghi R., *Mohankumar R., *Rozenberg D., *Coolens C., *Chan R., *Veit-Haibach P., *,**Wong AK.

*Joint Department of Medical Imaging, University Health Network, Toronto, ON, Canada
**Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

INTRODUCTION: Cartilage loss is a hallmark of knee osteoarthritis (OA). However, non-overweight individuals with knee pain have been poorly studied. It is known that bone marrow lesions (BMLs) are associated with pain regardless of a weight-bearing component. BMLs have been shown to have higher perfusion, especially in later stages of disease. Knee pain has also been associated with pro-inflammatory cytokines, originating from within fat. At the knee, the infrapatellar fat pad (IPFP) is a potential contributor to inflammation and pain. While studies have shown the relationship between pain and synovitis on MRI, little is known about how perfusion properties of the subchondral bone, synovium and IPFP each contribute to knee pain and physical function.
It is hypothesized that higher blood flow within the subchondral bone, synovium, and infrapatellar fat pad is associated with worse knee-specific pain.

METHODS: Women 50 to 85 years of age, with BMI <= 25 kg/m^2, and varying degrees of knee pain (as assessed by the Knee OA Outcome Score) were recruited as a convenience sample (N=48). Those with rheumatoid arthritis, existent joint replacements, or contraindications to MR imaging were excluded. T1-weighted, sagittal dynamic contrast enhanced (DCE)-MR images were obtained with Gadolinium (Gd) injection. BMLs, cysts and chondral lesions if present, were identified by musculoskeletal radiologists and manually segmented (N=10). The synovia were also manually segmented into suprapatellar, intercondylar (notch), and infrapatellar (fat pad) regions. Pharmacokinetics of Gadolinium arrival was calculated using the Toft's model, yielding a fluid transfer K^{trans} (min^{-1}) coefficient which reflects the perfusion of blood to the region. A 60 second area under the time-to-signal-intensity curve constant $iAUC_{60}$ (units: mmol.kg^{-1}second), a model free indicator of permeability of blood to the region, was also obtained. **Statistical analysis**: K^{trans} and $iAUC_{60}$ measured at each region were related to KOOS pain reported daily or always (dichotomized as ≥ 3) using a binary logistic regression model, reporting odds ratios (OR) and 95% confidence intervals per standard deviation (sd) higher K^{trans} and $iAUC_{60}$ values. All models accounted for age, body mass index and pain medication(s).

RESULTS: Among 48 participants, mean age was 61.4±8.7 years and BMI was 22.50±3.10 kg/m^2.
IPFP K^{trans} (perfusion) (mean 0.012 min^{-1}, min 0.000: max 0.092, SD: ±0.017 min^{-1}) was associated with an increased odd for having knee pain (OR: 4.65 (1.21, 17.93)). Similarly, $iAUC_{60}$ (mean 0.013, min 0.000: max 0.106, SD: ±0.023) was also associated with an increased odd for having knee pain (OR: 3.96 (1.29, 12.18). Suprapatellar $iAUC_{60}$ and intercondylar K^{trans} showed only a marginally significant association with knee pain (OR: 2.12 (0.92, 4.86)) and OR (2.10, (0.95, 4.66)). K^{trans} values (mean 0.010, min 0.000: max 0.150, sd±0.025) in subchondral bone without BMLs and $iAUC_{60}$ values (mean 9.71x10^{-4}, min 8.41x10^{-6}: max 0.005, SD 0.001) did not show a significant association with knee pain (OR: 0.79, (0.31, 2.03)), (OR: 1.26 (0.64, 2.50)). Among ten participants with notable abnormalities, the K^{trans} signal (0.048±0.055) within the BMLs, cysts or chondral lesions were larger than the mean K^{trans} within each of the IPFP (0.012±0.017), intercondylar (0.013±0.012) or suprapatellar (0.016±0.021) values ($p<0.001$).

CONCLUSIONS: Greater perfusion and permeability within the IPFP was associated with reporting of daily or constant knee pain. These perfusion indicators within the IPFP region of the joint suggest an inflammatory component to pain in knee OA.

ACKNOWLEDGMENT: Bart Schraa from Siemens-Healthineers is thanked for his assistance. CORRESPONDENCE ADDRESS: andy.wong@uhnresearch.ca

THE ASSOCIATION BETWEEN HIP PAIN AND RADIOGRAPHIC HIP OSTEOARTHRITIS IN THE CHECK COHORT

*Rondas G.A.M., *, **Macri E.M., ***Oei E.H.G., *, **Bierma-Zeinstra S.M.A., *Runhaar J.

*Erasmus MC University Medical Center Rotterdam, Department of General Practice, Rotterdam, the Netherlands.
**Erasmus MC University Medical Center Rotterdam, Department of Orthopedics and Sports Medicine, Rotterdam, the Netherlands
***Erasmus MC University Medical Center Rotterdam, Department of Radiology & Nuclear Medicine, Rotterdam, the Netherlands

INTRODUCTION: Hip pain is the most reported symptom in patients with hip OA and is associated with functional disability, participation restriction and loss of independence, which pushes patients to seek medical care. Hip pain is therefore a key feature in classification and diagnostic criteria for hip OA. In addition to clinical features, radiography is used to determine radiographic OA severity. Although the association between pain and radiographic features has been examined thoroughly in knee specific research, the association between hip pain and radiographic hip OA remains uncertain, for research examining this association is scarce and inconsistent.

OBJECTIVE: To assess whether the presence of hip pain is associated with an increased risk of radiographic hip OA in patients presenting in primary care, using the baseline cross-sectional data of the Cohort Hip and Cohort Knee (CHECK) study.

METHODS: Participants were eligible for CHECK when aged between 45-65, experienced pain in the hip and/or knee and hadn't had a prior consultation or were within 6 months of their first consultation to the general practitioner. All hips were assigned to a self-reported painful or pain-free group. Weight-bearing antero-posterior (AP) radiographs of the pelvis were used to assign KLG. Definite and early radiographic hip OA were defined as KLG ≥2 and KLG ≥1 respectively. The prevalence of definite and early hip OA was assessed in the full study sample and assigned subgroups of painful and pain-free hips. The association between hip pain and definite and early radiographic hip OA was assessed using general estimating equations (GEE), to account for the two observations (hips) within each individual.

RESULTS: KLG scores were missing in 22 hips at baseline, resulting in 1982 hips (991 participants) included in this study. Mean age was 56 years and 79% of participants were women. The overall prevalence of definite radiographic hip OA was 11.0% (218/1982), while the prevalence in painful and pain-free hips was 13.3% (105/789) and 9.5% (113/1193) respectively. The prevalence of early radiographic hip OA in all included hips was 35.3% (700/1982), 41.2% (325/789) in painful hips and 31.4% (375/1193) in pain-free hips. Compared to pain-free hips, the odds ratio (OR) of painful hips for having definite radiographic hip OA was 1.51 (95% CI [1.16, 1.98], p=0.003). The OR of painful hips for having early radiographic hip OA was 1.47 (95% CI [1.24, 1.75], p<0.001).

CONCLUSION: In patients presenting in primary care, hip pain was significantly associated with both definite and early radiographic hip OA. However, the OR and the difference in hip OA prevalence were modest. Therefore, in patients presenting with hip pain in primary care, suspected for hip OA, general practitioners should be confident to clinically diagnose hip OA and start treatment without radiographic imaging, since the additional diagnostic value of radiographic evidence is limited.

SPONSOR: none
DISCLOSURE STATEMENT: none
CORRESPONDENCE ADDRESS: 425635gr@student.eur.nl

DOMAIN ADAPATATION ON OAI DATASET FOR UNSUPERVISED SEGMENTATION OF BONE AND CARTILAGE

*Felfeliyan B., **Hareendranathan A., *Kuntze G., **Jaremko J., *,***Ronsky J.

*McCaig Institute for Bone and Joint Health University of Calgary, Calgary, AB, Canada
**Department of Radiology & Diagnostic Imaging, University of Alberta, Edmonton, AB, Canada
***Mechanical and Manufacturing Engineering, University of Calgary, Calgary, AB, Canada

INTRODUCTION: Bone and cartilage segmentation from MR images is a crucial step toward objective evaluation of OA. The development of new deep learning techniques have enabled automatic bone and cartilage segmentation. However, most deep learning models are trained on a single medical image dataset (annotated by an expert) and may fail to perform well on the data collected from different scanners or acquisition protocols. Adapting these models for new datasets requires expert-labeled annotations from human readers with medical expertise, which is time-consuming and expensive. Therefore, it is becoming increasingly important to develop a deep learning pipeline broadly applicable to unseen or unannotated data. Unsupervised Domain Adaptation (UDA) and domain generation techniques enable an existing model to adapt to multiple datasets. Cycle consistent GAN (CycleGAN) is a popular UDA method, which translates one dataset (source data) to another dataset (target data) by changing the data distribution of the source data.
The availability of annotated MRI datasets in the Osteoarthritis Initiative (OAI) presents a great opportunity to train one network on the labeled dataset (target data) and apply it to unseen data (source data) from different scanners.

OBJECTIVE: To perform automated segmentation of bone and cartilage on unseen source MRI data, based on unsupervised domain adaptation from training MRI sequences from a large publicly available dataset of knee MRI (OAI).

METHODS: This study requires two sets of data: 1) the source data, 2) target data. The source data consisted of a T1 MRI dataset collected using a 3T 750 GE scanner (Ethics ID #REB15-0554) at the University of Calgary. For the target dataset, 20 3D sagittal DESS scans were randomly selected from the OAI dataset with the segmentation labels from OAI-ZIB. These MRI slices from both datasets are the CycleGAN input. The CycleGAN is constructed from two generators and two discriminators. During the training, the first generator ($G_{T1\rightarrow DESS}$) receives data from the source dataset and translates it to the target domain (generates Pseudo-DESS images). The first discriminator (D_{DESS}) receives the Pseudo-DESS images and learns to detect whether its input is a real DESS image. The feedback from the D_{DESS} goes to $G_{T1\rightarrow DESS}$, and enforces the mapping from T1 to DESS. The same procedures are performed in the second generator ($D_{DESS\rightarrow T1}$), and the second discriminator (D_{T1}). The configuration of these four components reinforces the reverse mapping between two MRI modalities. As a result, structural information in the source image (T1) is preserved in the generated target image (pseudo DESS) Following training of the CycleGAN, the $G_{T1\rightarrow DESS}$ is used to translate T1 to Pseudo DESS. The result is forwarded to a convolutional neural network (CNN) trained on DESS sequences (Target) for segmentation of bone and cartilage (Figure 4). The dice similarity coefficient was used to compare the result of the segmentation between the predicted output and the ground truth.

RESULTS: For a pseudo-DESS scan with 204 slices, the segmentation model was successful in finding the location of bone or cartilage for 92% of the slices. The dice score for the detected bones is 0.83 and for cartilages is 0.64, indicating high agreement between the unsupervised segmentation and the labels.

CONCLUSION: Using CycleGAN, we proposed and validated a framework for knee MRI images for unsupervised segmentation. By translating images using our method, automatic and accurate segmentation of new knee MRI data was enabled without re-training the segmentation model.
This approach could lead to the ability to perform automated, unsupervised segmentation and analysis of MRI using different sequences from different institutions. Domain translation of sequences would lead to generalizability of automated MRI analysis in OA. In combination with the availability of the OAI dataset and annotations, the development of domain adaptation methods provides us with the opportunity to apply deep learning models trained on the OAI dataset to clinical sequences as well as new research sequences.

Fig 3: Unsupervised segmentation using domain translation pipeline. Source Domain (no ground truth; clinical MRI scans, eg. PDFS), Pseudo-Target Domain (MRI sequence matching research-protocol scan, eg. 3D DESS)

SPONSOR: Alberta Innovates, AHS Chair in Diagnostic Imaging, Medical Imaging Consultants, CIHR.
CORRESPONDENCE ADDRESS: banafshe.felfeliyan@ucalgary.ca

APPLICATION OF SUPER-RESOLUTION DEEP LEARNING FOR CLINICAL CONE-BEAM CT OF ANKLE JOINT

*Rytky S.J.O, *Finnilä M.A.J., **,*Karhula S.S., **Valkealahti M., **Lehenkari P., ***Joukainen A., ***Kröger H, ****Korhonen R.K., *****,*Tiulpin A., *,**Niinimäki J., *,**Saarakkala S.

*University of Oulu, Oulu, Finland
**Oulu University Hospital, Oulu, Finland
***Kuopio University Hospital, Kuopio, Finland
****University of Eastern Finland, Kuopio, Finland
*****Aalto University, Espoo, Finland

INTRODUCTION: Weight-bearing cone-beam CT (CBCT) allows examining joint behavior in the physiological context. Especially, trabecular bone is metabolically active, making it an attractive target for developing OA imaging biomarkers. However, the spatial resolution of current clinical CBCT systems is not sufficient to visualize subtle microstructural changes in the subchondral bone. Modern deep-learning-based super-resolution methods can computationally increase the spatial resolution when trained both with low- and high-resolution images. In principle, these methods could be suitable also for the prediction of bone microstructural features from clinical CBCT. In this study, we apply the super-resolution method to a clinical CBCT scan of the ankle joint, i.e., one of the most often imaged joints with CBCT. Osteochondral ex vivo samples, imaged with µCT, were used to train the model.

OBJECTIVE: Use ex vivo µCT data to enhance the clinical CBCT image quality with deep learning super-resolution methods.

METHODS: Tissue blocks (n = 4) were extracted from one TKA patient and two cadavers, imaged with preclinical µCT (Bruker Skyscan 1176; 80kV, 125µA, 26.7 µm voxel size) and reconstructed. Furthermore, one clinical weight-bearing CBCT scan of an ankle was extracted from the hospital archive (Planmed Verity; 96kV, 8mA, 400 µm voxel size), and a volume-of-interest (VOI) of 5x6x7cm was selected (Figure 1d). Training data was created by scaling the µCT images to a resolution of 400µm and 100µm with an antialiasing filter applied (Figure 1a-b). Deep convolutional neural networks were trained to magnify the µCT images to the higher resolution. Both 2D and 3D models were used with combinations of the mean-squared error (MSE), mean absolute error (MAE), total variation (TV), structure similarity index (SSIM) and perceptual loss (PL) with spatial and intensity-based augmentations applied. Out-of-fold SSIM scores were calculated against the µCT ground truth (radius = 7). Each experiment was conducted with six random initializations to estimate the 95% confidence intervals. Finally, the ankle VOI was passed through the networks and visually evaluated.

RESULTS: The highest out-of-fold performance was achieved by the 2D model with MSE and TV loss (SSIM = 0.69 ± 0.01). Despite lower results (SSIM = 0.23 ± 0.01), the respective 3D model achieved very good image quality on the ankle VOI. The 2D model with MAE, PL and TV losses yielded the best visual quality (SSIM = 0.67 ± 0.005, Figure 1c, 1e). The use of SSIM loss did not improve the results.

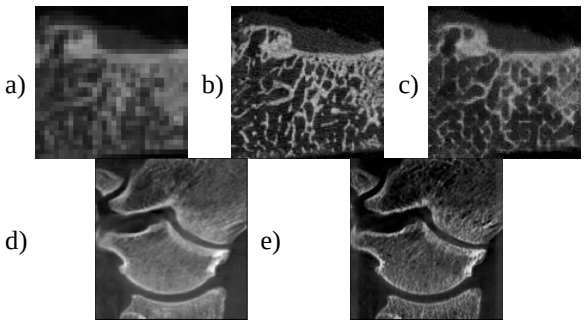


Figure 1. Input (a) and target (b) images from the preclinical µCT training set. Out-of-fold prediction of the 2D model (MAE+TV+PL loss) (c). Example of the clinical ankle CBCT (d), and the 2D model output (e).

CONCLUSION: Visually, the best 2D models produced a plausible resolution increase for both the training set (Figure 1a-c) and the test set (Figure 1d-e). The 2D models yielded the highest out-of-fold SSIM against the µCT ground truth. However, the perceived quality was also high with the 3D model predictions. Interestingly, the method generalizes well despite the distinct image quality differences between CBCT and µCT, and the small sample size. In the future, the method should be validated in a larger dataset and the results on clinical CBCT should be quantified with suitable metrics, e.g., bone morphometrics.

SPONSOR: Instrumentarium Science Foundation (No. 200058)
DISCLOSURE STATEMENT: None
ACKNOWLEDGMENTS: CSC—IT Center for Science, Espoo, Finland
CORRESPONDENCE ADDRESS: santeri.rytky@oulu.fi

INTERPRETABLE DEEP LEARNING FRAMEWORK FOR KNEE OSTEOARTHRITIS STRUCTURAL PROGNOSIS PREDICTION

*Nguyen H.H., *,**,***,****Tiulpin A., *,****Saarakkala S.

*Research unit of Medical Imaging, Technology, and Physics, University of Oulu, Finland
**Aalto University, Finland
***Ailean Technologies Oy, Finland
****Department of Diagnostic Radiology, Oulu University Hospital, Finland

INTRODUCTION: Recent studies have focused on developing deep learning (DL) methods to predict the current severity of knee OA or the progression of the disease within a pre-defined time period. However, in progression prediction the use of fixed time frame and the interpretability of DL models have been limitations so far. The ability to predict knee OA progression at multiple time points in the future, instead within a single time period, would be a significant step for DL-based prediction models towards their implementation in clinical practice.

OBJECTIVE: To develop an interpretable DL framework to simultaneously predict the progression of knee OA structural changes in terms of KL grade at multiple time points in the future from a knee plain radiograph and clinical variables.

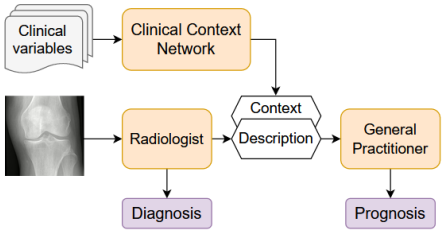


Figure 5: Our framework

METHODS: We used the data from the OAI cohort with 4,796 participants and prepared the training dataset as follows. Firstly, we generated data entries across the follow-up visits (baseline to 72-month), each of which included a plain radiograph of a single knee and clinical data (age, BMI, injury, sex, surgery, and total WOMAC) of a participant. Our prediction targets were KL grades at the present and the next 8 years at all available OAI follow-ups. After filtering out invalid cases with decreasing KL grades, we obtained 38,363 data entries in total. In our experiments, we followed an out-of-hospital cross-validation strategy in which we iteratively trained our models on samples acquired at 4 hospitals and tested on data from the left-out one. We assessed the performance and calibration of our models via balanced accuracy (BA) and expected calibration error (ECE), respectively.
Our framework was inspired by a clinical decision-making process illustrated in Figure 1, which we implemented via three transformer-like neural networks, passing information to each other. As such, a "clinical context network" and a "radiologist network" extract context and visual embeddings from clinical variables and radiographs, respectively. Subsequently, while the "radiologist network" predicts the current KL grade, a third network, acting as a "general practitioner", relies on the given context and visual description to predict the future KL grades.

RESULTS: When predicting the course of the disease in the near future within 4 years after the initial time point, our models achieved reliable performances with BAs of 63.3±0.02%, 55.8±0.03%, 50.8±0.3%, and 44.5±0.02%, and ECEs of 3.6±0.02%, 4.1±0.03, 6.0±0.2%, and 6.1±0.1%, respectively. As expected, their performances dropped significantly to BAs of 28.5±0.07% and 26.4±0.1% when predicting the structural prognosis at the 6- and 8-year time frames. Importantly, besides just predicting the prognosis in the future, our models also provide multi-modal interpretations (feature importances) for both imaging and non-imaging input data (Figure 2).

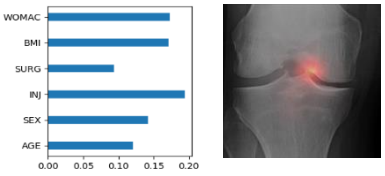


Figure 2: (Left) Clinical variables' attention. (Right) Radiograph's attention maps.

CONCLUSION: We present the interpretable, end-to-end DL framework that can learn from multi-modality data to predict the knee OA structural progression at multiple future time periods. Our results suggest that DL models combining plain radiography with clinical data can predict the knee OA structural prognosis within the near future (1-4 years). However, it seems that predictions to the far future (6-8 years) is more challenging for these DL models when using only plain radiography and clinical data.

SPONSOR:
DISCLOSURE STATEMENT: A.T. is a co-founder of Ailean Technologies Oy, a medical technology company providing consulting services.
ACKNOWLEDGMENT: The OAI participants, sites, investigators, funders and coordinating center.
CORRESPONDENCE ADDRESS: huy.nguyen@oulu.fi

VIRTUAL BONE AGING: CAN WE PREDICT 48 MONTHS BONE SURFACE CHANGES?

*Calivà F., *Morales M.A., *Majumdar S., *Pedoia V.

*Center for Intelligent Imaging (CI2), University of California, San Francisco, CA, USA

INTRODUCTION: Knee OA bone shape changes are an appealing outcome target for clinical trials, and patient-personalized bone changes trajectory prediction could be of great interest in understanding the impact of specific intervention strategies.

OBJECTIVE: To use deep learning (DL) to predict knee bone surface changing in a 48-month time frame, in subjects with and without OA.

METHODS: This study use data from the OAI. 3D-DESS MRI were used to segment the knee bones using a V-Net model previously developed and validated. From segmentation masks, shape features were encoded in 2D spherical maps, which have previously proven powerful in diagnosing and predicting future OA. We used the femur bone spherical maps from three time points (baseline, 12 months, and 24 months) as inputs to a model tasked with predicting the longitudinal bone surface changes at the 72-month time point. 4133 subjects were selected and split into 2855/598/680 samples per train/val/test set, controlling for BMI (27.41±4.53/27.05±4.35/27.26±4.07 kg/m2), age (62.37±9.01/62.55±9.11/60.64±8.27), and sex (415-569/84-122/114-112 male/female) statistical independence across the train/val/test splits. To achieve our goal, we implemented three solutions. (1) involved a 2D modified V-Net (baseline model). This model was given the concatenation of the spherical bone encoding of the three initial time points as an input, and it produced a prediction of the bone shape at a subsequent time point (72 months) (2), we corrupted the input spherical maps with white noise (mean=0, var=0.03), and utilized the 2D V-Net as a denoising autoencoder. In (3) namely Δshape, we guided the network with an explicit supervision for the bone shape changes. We engineered the network target as the difference between the bone spherical maps at the 24- and 72-month time points. All the V-Nets consisted of 5 levels with 3, 4, 3, 4, 2 convolutions per level. Parameters were tuned by minimizing a mean-absolute error (MAE) and structural similarity index metric (SSIM) loss, rescaled using a 6.7 factor. Batch size was set to 8, a 0.05 dropout rate was implemented. Adam optimizer with learning rate 1×10-5 was adopted.

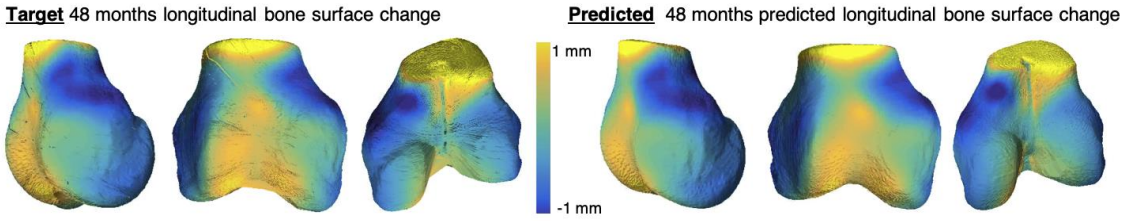
Table 1 model results

	All (N=680)		OA (N=39)		Not-OA (N=641)	
Model	MAE	SSIM	MAE	SSIM	MAE	SSIM
Enc-Dec	1.409 (0.653)	0.922 (0.021)	1.496 (0.458)	0.920 (0.019)	1.404 (0.662)	0.922 (0.021)
Denoiser	1.653 (0.648)	0.926 (0.022)	1.768 (0.517)	0.924 (0.019)	1.646 (0.655)	0.926 (0.022)
Δ Shape	1.076 (0.630)	0.912 (0.024)	1.080 (0.444)	0.911 (0.021)	1.075 (0.639)	0.912 (0.024)

RESULTS: Bone segmentation was highly accurate (Dice score=0.98±3.2). Average errors computed in the test set in terms of MAE on the surface (mm) and SSIM are reported in Table 1 for the whole population and stratified by OA status. Δshape model yielded the best performances. A visual example comparing imaging derived and predicted bone shape changes is reported in Figure 1.

CONCLUSIONS: We predicted femoral bone remodeling in a cohort of patients with OA and control groups 48 months ahead of time. We observed the most promising solution was Δshape, which supervised the network over the actual bone shape changes. While this is just a pilot exploration, we demonstrated DL models can capture and predict trajectory of local changes in bone surface shape over time.

SPONSOR: NIH-NIAMS R00AR070902 (VP)
DICLOSURE STATEMENT: No Conflict of Interest to disclose
CORRESPONDENCE ADDRESS: francesco.caliva@ucsf.edu



KNEE JOINT DISTRACTION RESULTS IN MRI CARTILAGE THICKNESS INCREASE UP TO TEN YEARS AFTER TREATMENT

*Jansen M.P., *Mastbergen S.C., **Turmezei T.D., ***MacKay J.W., *Lafeber F.P.J.G.

*University Medical Center Utrecht, Utrecht, The Netherlands
**Norfolk & Norwich University Hospital, Norwich, UK & University of East Anglia, Norwich, UK
***University of East Anglia, Norwich, UK & University of Cambridge, Cambridge, UK

INTRODUCTION: Knee joint distraction (KJD) is a surgical joint-preserving treatment option for younger (age <65 years) knee OA patients. It has been shown to provide significant clinical improvement for up to nine years after treatment. Furthermore, both radiographs and MRI scans have previously shown cartilage regeneration activity, especially in the first two years after treatment. However, MRIs have not been evaluated more than five years after this treatment.

OBJECTIVE: To evaluate MRI cartilage thickness up to ten years after KJD treatment.

METHODS: Patients (n=20) with end-stage knee OA, indicated for TKA but <60 years old, were treated with KJD. 3T MRIs with 3D spoiled gradient recalled imaging sequence with fat suppression (SPGR-fs) were acquired before and one, two, five, seven and ten years after surgical treatment. Stradview v6.0 was used for semi-automatic cartilage segmentation; wxRegSurf v18 was used for surface registration. MATLAB R2020a and the SurfStat MATLAB package were used for data analysis and visualization. For changes over time, a linear mixed model was used. Statistical significance was calculated with statistical parametric mapping; a p-value <0.05 was considered statistically significant. Patients were separated in two groups based on whether their most affected compartment (MAC) was the medial or lateral compartment.

RESULTS: One and two years after treatment the cartilage in the MAC weight-bearing region was on average thicker than before treatment. While from five years after treatment the cartilage thickness gradually decreased, even at ten years the medial cartilage thickness seemed slightly higher than before treatment. The medial cartilage thickness increase in patients with a medial MAC (n=18), which was up to 0.5 mm after one year and 0.6 mm after two years, was partly statistically significant at both these time points (Figure 1). Surprisingly, long-term results showed areas of the lateral (less affected) compartment were significantly thicker, up to 0.7 mm, compared to pre-treatment in both the femur and tibia compared to baseline.

CONCLUSION: KJD treatment results in significant short-term cartilage regeneration in the most affected compartment. While after two years this initial gain in cartilage thickness is gradually lost, likely as a result of natural progression, even ten years after treatment the cartilage is thicker than before treatment. In the less affected compartment, a delayed cartilage response seems to take place, with significantly increased cartilage thickness in the long term. In conclusion, in these young OA patients indicated for TKA, KJD results in femoral and tibial cartilaginous tissue regeneration both short- and long-term and in both sides of the joint.

SPONSOR: ReumaNederland (Dutch Arthritis Society, project number LLP-9).
DICLOSURE STATEMENT: T.D. Turmezei is on the IWOAI 2021 scientific committee.
CORRESPONDENCE ADDRESS: m.p.jansen-36@umcutrecht.nl

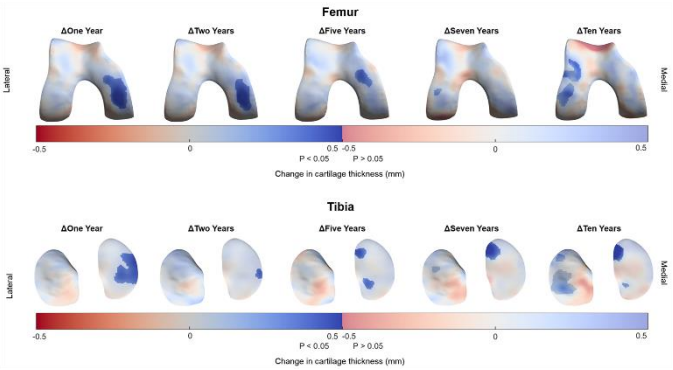


Figure 1: The change in cartilage thickness compared to baseline, for all patients whose medial compartment was the most affected, after one (n=18), two (n=18), five (n=15), seven (n=11) and ten (n=7) years after treatment with knee joint distraction. Statistically significant changes are indicated by the darker colour map (p<0.05) while non-significant areas are shown with faded colours (p>0.05). Blue indicates an increase and red a decrease in cartilage thickness compared to baseline. Results are viewed on average right femur and tibia articular cartilage surfaces.

FOCAL CHANGES IN T_2 RELAXATION TIME OF LOADED CADAVER FEMORAL CARTILAGE

*Bzowey N.B., **Black M.S., *Cui L., *Elebute I., *Thorson C.S., *Martel M.M., *McWalter E.J.

*Department of Mechanical and Biomedical Engineering, University of Saskatchewan, Saskatoon, Canada
**Department of Radiology, Stanford University, Stanford, CA, USA

INTRODUCTION: T_2 relaxation times of articular cartilage increase with OA[1] and are associated with collagen organization and water content[2,3]. Most qMRI studies of cartilage acquire data in the unloaded knee but it has been shown that T_2 relaxation times decrease in some regions when the knee is loaded[4]. However, regional assessments do not isolate where, within the region, the changes occur.

OBJECTIVE: To identify focal changes in T_2 relaxation times in loaded cadaver femoral cartilage.

METHODS: Six cadaver knee specimens (70.3±9.3 years; no history of injury/surgery; no full thickness cartilage defects) were scanned using a quantitative double-echo in steady-state (qDESS) sequence at 3T (Magnetom Skyra, Siemens, Germany). Imaging parameters were TR/TE1/TE2 = 16.5, 3.75, 29.3 ms; matrix size: 256 x 256; field of view: 160 mm x 160 mm; slice thickness: 3mm. Specimens were scanned unloaded and then under an axial load of approximately 800 N using a custom MRI-safe loading rig. There was a 110-minute delay between the application of the load and the image acquisition to account for cartilage stress relaxation. Femoral cartilage was manually segmented (Analyze 14.0, AnalyzeDirect, USA) and T_2 relaxation time maps were generated[5]. Projection maps were created, and clusters of increased (positive) and decreased (negative) T_2 relaxation time were identified[6,7]. A cluster was defined as groups of contiguous pixels for which 1) T_2 relaxation time was greater than two times the standard deviation of the difference between the unloaded and loaded maps, and 2) area was greater than 1% of the total projection map area. The area of the cartilage plate covered by positive clusters was ΔT2%CA+ and by negative clusters was ΔT2%CA-. We determined if ΔT2%CA+ and ΔT2%CA- were different from zero using a one-sample Wilcoxon Signed Rank test (SPSS, IBM, USA). Descriptive statistics of clusters for the following regions were also reported: lateral-anterior (LA), lateral-central (LC), lateral-posterior (LP), medial-anterior (MA), medial-central (MC), and medial-posterior (MP).

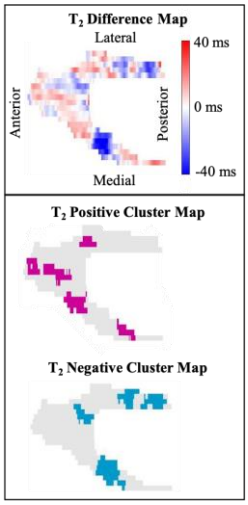


Figure 1: Sample maps from specimen 1

RESULTS: Clusters of increased and decreased T_2 relaxation time with load were observed (Figure 1) and the areas were significantly different from zero (ΔT2%CA+: median=7 %, range=0 to 17%, p=0.043; ΔT2%CA-: median =17%, range=8 to 18%, p=0.028). The presence of clusters varied by region, with most appearing in the medial and lateral central regions (Figure 2).

CONCLUSION: The negative clusters covered more area than the positive clusters which aligns with previous results that found a net decrease in T_2 relaxation time with load within this region[4]. Most of the clusters appeared in the medial and lateral central compartments which is not surprising since this is the region that was directly loaded. The presence of both positive and negative clusters may be indicative of unequal distribution of load throughout the cartilage. Future work will examine the relationship between the clusters and specific loaded areas. In conclusion, both positive and negative focal changes in T_2 relaxation time are present in loaded cartilage and appear predominantly in loaded regions.

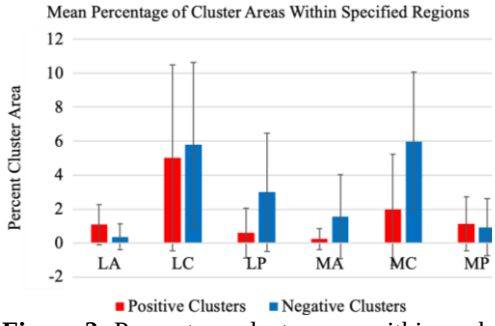


Figure 2: Percentage cluster area within each region. Most clusters are within the lateral and medial central regions

SPONSOR: NSERC, Mitacs, U of Saskatchewan
DISCLOSURE: We receive research support from Siemens (EJM, LC) and GE (MB).
CORRESPONDANCE: natasha.bzowey@usask.ca
REFERENCES: 1. Dunn, Radiology 2004; 2. Raya, ISMRM 2017; 3. Nieminen, MRM 2001; 4. Subburaj, JOR 2012; 5. Sveinsson, MRI 2017 6. Monu, OA&C 2016; 7. Black, Stanford Thesis 2019.

FUNCTONAL IMAGING OF CARTILAGE BY SERIAL T1ρ MAPPING - DIFFERENT LOADING REGIMES IN THE ASSESSMENT OF TISSUE FUNCTIONALITY

*,**Nebelung S, **Zwingenberger K, *Schock J, **Truhn D

*Department of Diagnostic and Interventional Radiology, University Hospital Düsseldorf, Düsseldorf, Germany
**Department of Diagnostic and Interventional Radiology, Aachen University Hospital, Aachen, Germany

INTRODUCTION: Increasing evidence suggests that functional MRI of articular cartilage, i.e., imaging the tissue's response to loading, may help assess the tissue's functionality in health and disease. When subject to loading, (histologically) intact cartilage displays other response-to-loading patterns than (histologically) degenerative cartilage.

OBJECTIVE: The aim of this study was to assess the responses to static and dynamic loading (as compared to no loading) of histologically intact human articular cartilage samples under pressure-controlled *in-situ* loading by use of serial T1ρ mapping. T1ρ was selected as the quantitative MRI parameter of interest as it is exquisitely mechanosensitive and allows structural and compositional intra-tissue adaptations beyond mere morphology. Our hypothesis was that static and dynamic loading bring about significantly different spatial and temporal changes in T1ρ that were reflective of intra-tissue adaptations.

METHODS: 47 cartilage samples were harvested from the medial femoral condyles of patients undergoing total knee arthroplasty (pooled from 24 patients [15 women, 9 men; mean age 67.4 years [range, 50-82 years]), prepared to standard thickness, and placed in a standard knee joint of a pressure-controlled whole knee-joint compressive loading device (Fig. 1a). Individually, the samples' responses to static (i.e., constant at 4 bar [84 % body weight]), dynamic (i.e., cyclically alternating between 0 and 4 bar at 0.5 Hz), and no loading (i.e., free-swelling conditions) were assessed before (δ_0), and after 30 min (δ_1) and 60 min (δ_2) of loading (Fig. 1b). Proton density-weighted images and T1ρ maps were serially acquired on a clinical 3.0T MRI scanner (Achieva, Philips) (Fig. 1c, d)and relative changes in T1ρ (Δ_1, Δ_2) were determined for the entire sample and the superficial and deep layers as a function of time. MRI responses to loading were reference to histologic (i.e., Mankin score) and biomechanical reference measures (i.e., tangent stiffness).

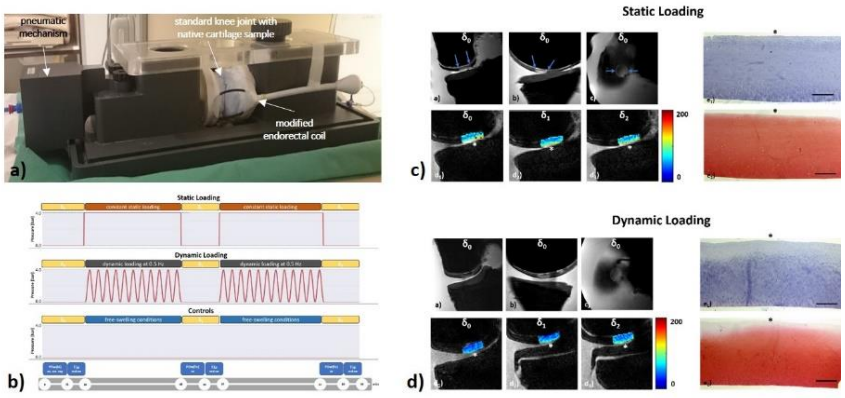


Figure 1: a) Details of MRI-compatible pressure-controlled whole-knee joint loading device. b) Detailed timelines of loading regimens and serially acquired MRI sequences. Unloaded reference configuration (δ_0), following 30 min (δ_1) and 60 min (δ_2) of static, dynamic or no loading. c, d) Representative chondral samples' responses to static (c) and dynamic (d) loading and histologic cross-referencing as assessed by morphologic PD-weighted images of the medial femorotibial compartment and color-coded T1ρ maps at δ_0, δ_1, and δ_2.

RESULTS: Relatively homogeneous histologic, biomechanical, and T1ρ sample characteristics were found at δ_0. Under loading, relative increases in T1ρ were significant after dynamic loading (Δ_1=10.3±17.0 %, Δ_2= 21.6±21.8 %, p=0.002), while relative increases in T1ρ after static or no loading were only moderate and not significant.

CONCLUSION: This experimental *in-situ* study suggests that dynamic loading and T1ρ mapping are characterized by a larger dynamic range than corresponding static loading to assess cartilage functionality in patients based on advanced MRI techniques complemented by biomechanical loading.

SPONSOR: Deutsche Forschungsgemeinschaft (DFG, Grant Nr. NE 2136/3-1).
DISCLOSURE STATEMENT and ACKNOWLEDGEMENT: None.
CORRESPONDENCE ADDRESS: sven.nebelung@med.uni-duesseldorf.de

COMPREHENSIVE SERIAL T2 MAPPING TO EVALUATE AND MONITOR POST-TRAUMATIC HUMAN CARTILAGE DEGENERATION

*Huppertz M.S., **Schock J., *Truhn D., **Nebelung S.

*Department of Diagnostic and Interventional Radiology, Aachen University Hospital, Aachen, Germany
**Department of Diagnostic and Interventional Radiology, Medical Faculty, University Düsseldorf, Düsseldorf, Germany

INTRODUCTION: Supraphysiologic impaction injury as occurring in severe trauma predisposes joint surfaces to degenerative changes and posttraumatic osteoarthritis. Morphologic MRI as well as other current clinical-standard imaging modalities lack the ability to detect such injuries and monitor their progression to posttraumatic degeneration with oftentimes serious prognostic implications. Advanced MRI acquisition and post-processing techniques such as T2 mapping combined with texture feature analysis are promising complementary techniques that may fill this diagnostic gap, yet have not been thoroughly evaluated in controlled basic research contexts.

OBJECTIVE: Our hypotheses were that 1) variable impaction energies induce variable progressive posttraumatic degenerative changes in human articular cartilage and that 2) these changes are reflected qualitatively by T2 mapping features and quantitatively by associated descriptive statistics and texture features.

METHODS: Macroscopically intact human articular cartilage samples obtained from total joint replacements (n = 35, central lateral femoral condyle) were exposed to standardized injurious impaction with low (0.49 J, n = 14) or high (0.98 J, n = 14) energy levels using a custom-made drop tower device (**Fig. 1a**). Samples were imaged before (t_0), immediately after impaction (t_1), and after another 24 h (t_2) and 72 h (t_3) of standard cultivation using a T2 mapping (multi-spin echo) sequence and a clinical 3.0 T MRI scanner (Achieva, Philips) (**Fig. 1b**). While sample outlines were segmented manually, another two regions-of-interest, i.e., the superficial and deep sample layers, were determined automatically. Additionally, contrast, homogeneity, energy, and variance were quantified as textural features based on gray-level co-occurrence matrices. Longitudinal (time-related) differences were assessed by the Friedman test and cross-sectional (group-related) differences at individual time points by the Kruskal-Wallis test. Unimpacted controls (n = 7) and histology (**Fig. 1c**) served as references.

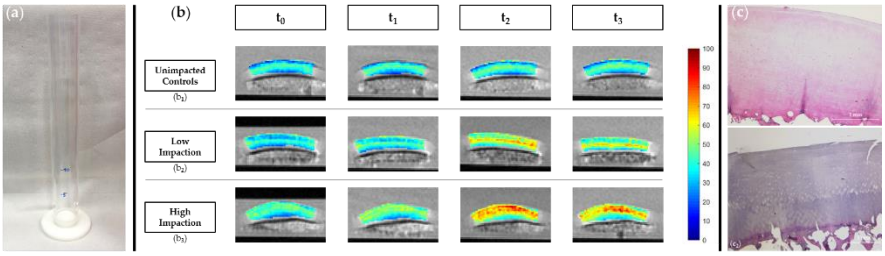


Figure 1: a) Custom-made drop-tower device for cartilage impaction. b) Representative T2 maps of cartilage samples following no impaction (controls, b_1), low (b_2), and high impaction (b_3). c) HE-stained histologic sections of an intact sample before (c_1) and after high-impaction injury (c_2).

RESULTS: Absolute T2 values ($p < 0.001$), contrast ($p < 0.001$), and variance ($p = 0.005$) were significantly increased as a function of impaction energy and time. Homogeneity ($p = 0.001$) and energy ($p = 0.006$) were significantly decreased. Changes in T2 involved the entire sample thickness, yet tended to be more pronounced in the superficial layer. Qualitative histologic reference assessment revealed only moderate surface disintegration and loss of proteoglycan staining-intensity in impacted cartilage samples.

CONCLUSION: The severity of post-traumatic cartilage degeneration is related to injurious impaction energy. T2 and texture features are sensitive diagnostic markers to detect and monitor traumatic impaction injuries of cartilage and associated post-traumatic degenerative changes and may be used to assess cartilage after trauma to identify “cartilage tissue at risk”.

SPONSOR: 1) Deutsche Forschungsgemeinschaft (DFG, Grant Nr. NE 2136/3-1), 2) START Program of the Faculty of Medicine, RWTH Aachen, Germany.
DICLOSURE STATEMENT and ACKNOWLEDGMENT: None.
CORRESPONDENCE ADDRESS: mhuppertz@ukaachen.de

POSTERS

EXPLORING SUBCHONDRAL BONE CHANGES MEASURED WITH CT AFTER JOINT DISTRACTION TREATMENT FOR END STAGE KNEE OSTEOARTHRITIS

*Jansen M.P., *Ooms A., **Turmezei T.D., ***MacKay J.W., *Mastbergen S.C., *Lafeber F.P.J.G.

*University Medical Center Utrecht, Utrecht, The Netherlands
**Norfolk & Norwich University Hospital, Norwich, UK & University of East Anglia, Norwich, UK
***University of East Anglia, Norwich, UK & University of Cambridge, Cambridge, UK

INTRODUCTION: Knee OA can cause significant bone changes, including cortical bone thickening and subchondral bone density decrease and a change in overall bone shape. Knee joint distraction (KJD) is a joint-preserving surgical treatment for younger (<65 years) knee OA patients that has been shown to reverse OA cartilage degradation. On radiographs, KJD patients showed a decrease in subchondral bone density and an increase in osteophyte formation after treatment. However, these bone changes have never been evaluated with a 3D imaging technique.

OBJECTIVE: To evaluate cortical bone thickness, subchondral trabecular bone density, and bone shape on CT scans before and one and two years after KJD treatment.

METHODS: In total 19 patients underwent a CT scan before KJD treatment and one and two years after treatment. Stradview v6.0 was used for semi-automatic segmentation of the tibia and femur from axial thin-slice (0.45mm) CT scans, from which cortical bone thickness (mm) and trabecular bone density (Hounsfield units, HU) were measured with an automated algorithm. Osteophytes were excluded from segmentation. Afterwards, wxRegSurf v18 was used for surface registration. MATLAB R2020a and the SurfStat MATLAB package were used for data analysis and visualization. Two-tailed F-tests were used to calculate changes over time. Statistical significance was calculated with statistical parametric mapping; a p-value <0.05 was considered statistically significant. Bone shape changes were explored visually from the 3D models using the vertex by vertex displacement of the femur and tibia between time points. Patients were separated into two groups based on whether their most affected compartment (MAC) was medial or lateral.

RESULTS: 16 Tibias and 15 femurs could be analyzed. For both the femur and tibia, before treatment the cortical bone was thicker and trabecular bone density higher in the MAC compared to the rest of the joint. After treatment, the MAC cortical thickness decreased, although this decrease of up to ~0.25 mm over the first year was not statistically significant. A decrease in trabecular bone density was seen throughout the entire joint (Figure 1), although significant only for small areas on mostly the medial side in patients with a medial MAC (n=14), where this decrease was up to ~80 HU over the first year. Lastly, bone shape analyses showed similar general shape changes for both sides of the joint and both the femur and tibia especially in the second year after treatment: the central areas of both compartments showed an outward change, while the outer ring of both compartments showed inward changes.

CONCLUSION: Especially in the MAC, KJD results in thinning of the subchondral cortical bone plate and decrease of subchondral bone density in the first year after treatment, which is sustained throughout the second year. For both parameters, the MAC became more similar to the LAC, indicating (partial) subchondral bone normalization in the most affected parts of the joint. The bone shape changes may indicate a reversal of typical OA changes, since femoral condyles became more convex and tibial condyles less concave. However, the inward difference that was seen on the outer edges may be a result of osteophyte-related changes as well, since this might have affected the bone segmentation.

SPONSOR: ReumaNederland (Dutch Arthritis Society, project number LLP-9).
DICLOSURE STATEMENT: T.D. Turmezei is on the IWOAI 2021 scientific committee.
CORRESPONDENCE ADDRESS: m.p.jansen-36@umcutrecht.nl

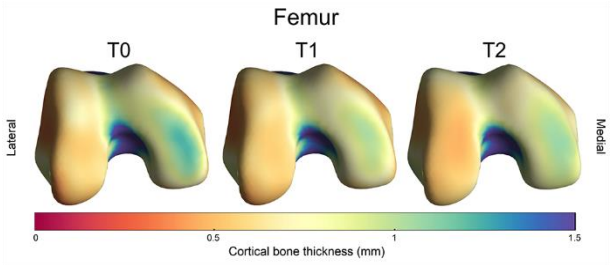


Figure 1: Femoral trabecular bone density of patients with predominantly medial compartmental osteoarthritis (n=14), before and one and two years after treatment with knee joint distraction.

INTER- AND INTRARATER RELIABILITY, AND DIAGNOSTIC TEST ACCURACY OF A KNEE OSTEOARTHRITIS MACHINE LEARNING DECISION-AID ALGORITHM IN A NEAR-CLINICAL SETTING

*Brejneboel M.W., *Hansen P., *Nybing J.D., *,****Bachmann R., *Ratjen U., **Hansen I.V., ***Axelsen M., ***Lundemann M., *Boesen M.

*Department of Radiology, Bispebjerg and Frederiksberg Hospital, Copenhagen, Denmark
**Department of Radiology, Herlev and Gentofte Hospital, Copenhagen, Denmark
***Radiobotics ApS, Copenhagen, Denmark
****Radiography, Department of Technology, University College Copenhagen, Denmark

INTRODUCTION: Overall increases in medical imaging, scarcity of radiologists, and concerns on interrater variability have sparked efforts to develop robust and time saving auto-reporting. Plain radiography remains mainstay in the diagnostic workup of knee osteoarthritis (KOA). The purpose of this study was to evaluate the agreement between a CE-marked machine learning (ML)-tool (RBknee™, Radiobotics ApS) for KOA decision-aid and various radiology professionals.

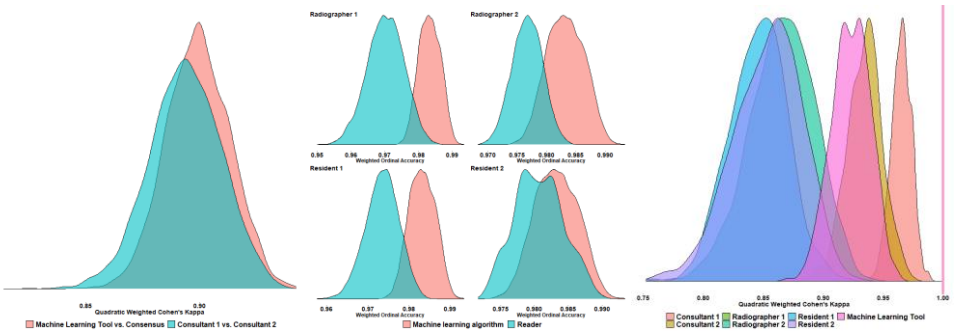
OBJECTIVES: 1) Does the ML-tool perform reliably compared to human readers with different levels of experience in musculoskeletal radiology? 2) What is the diagnostic test accuracy of the ML tool compared to that of a consensus between experienced musculoskeletal radiology consultants?

METHODS: In this retrospective, single-center, interrater reliability study, two musculoskeletal radiology consultants, two experienced reporting radiographers, two resident radiologists, and the ML-tool rated 50 consecutive posterior-anterior (Rosenberg's method) plain radiographs of the weight-bearing knee (n=100). Before rating began, a modified OARSI atlas on knee Kellgren-Lawrence (KL) grading was issued to all readers. Each reader independently rated each knee on the ordinal KL grade. Quadratic Weighted Cohen's kappa and Weighted Ordinal Accuracy (as proposed by Obuchowski) were calculated as a metrics for reliability and accuracy. Both used a quadratic penalty matrix. During computation, only one knee from each study was sampled to avoid intra-patient bias.

RESULTS: The ML-tool showed good-to-excellent interrater reliability with consultant consensus with kappa of 0.9 (CI95%: 0.86-0.94). Initial interrater agreement among musculoskeletal consultants was 0.89 (CI95%: 0.86-0.93). Similar results were found for the other readers.
Intrarater agreement for consultants were: 0.96 (CI95%: 0.94-0.98) and 0.94 (CI95%: 0.91-0.97). For the ML-tool: 1 (CI95%: 1-1).
Post-hoc, consultants arrived at consensus for 34 knees on which they had disagreed. Weighted Ordinal Accuracy for the ML-tool was 0.98 (CI95%: 0.98-0.99), for radiographers 0.97 (CI95%: 0.96-0.98) and 0.98 (CI95%: 0.97-0.98), for residents 0.97 (CI95%: 0.97-0.98) and 0.98 (CI95%: 0.97-0.99).

CONCLUSION: The machine learning tool performed reliably compared to the consultant radiologist consensus. Because this study used consecutive data from clinical practice we expect similar performance in a clinical setting. It is therefore likely that the machine learning tool could reliably assist radiologists in the grading and reporting of knee osteoarthritis on radiographs.

SPONSOR: EUROSTARS, grant: E! 12835 - X-AID
DICLOSURE STATEMENT: M. Axelsen and M. Lundemann are employees of Radiobotics ApS. M. Boesen is a member of the scientific advisory board at Radiobotics ApS.
CORRESPONDENCE ADDRESS: mathiaswbrejne@outlook.com



DEEP LEARNING APPROACH TO FULLY-AUTOMATED BONE MARROW LESION SEGMENTATION

*Gardea M., *Anwari V., *Usman R., *Liu S., *Wong AK.

*Joint Department of Medical Imaging, University Health Network, Toronto, ON, Canada

INTRODUCTION: Pang et al (2013) found that it took between 8 and 35 minutes to manually segment Bone Marrow Lesions (BMLs) in a pair of MR images. In standard of care, the test-retest imprecision and lack of sensitivity of semi-quantitative BML grading obscures its correlations to patient outcomes.

OBJECTIVE: We aimed to automate the segmentation of BMLs using a deep learning approach.

METHODS: Our team validated a custom-designed UNET neural network architecture to segment BMLs within MRI slices of the knee. A subset of sagittal double echo steady state MR images at 3.0T (0.5x0.5x1.0mm) from 100 randomly-selected participants (N=100) in the Osteoarthritis Initiative (OAI) who were noted by radiologists as having a BML as per WORMS, MOAKS or BLOKS scoring, was used. BMLs in these 16,000 MRI slices (160 slices per patient) were manually annotated with polygonal ROIs using the LabelMe annotation tool. Annotations were verified by trained musculoskeletal radiologists. This dataset excluded cysts from both training and testing samples since the study aimed to exclusively segment BMLs. Patients were then split into training (80) and test (20) subsets. Hyperparameter choices included the number of starting filters, dropout rate, learning rate, and a multiplicative factor of steps per epoch (steps per epoch is a value generated by dividing the sample size by the batch size). The gradient descent optimizer used was Adam, known for its strong performance in convolutional neural networks. Implemented as well was a dataset augmentation function. The augmentation created new images based on a list of potential transformations including rotation, translation, horizontal flipping, window/leveling, and zooming. Performance was measured by the 1-DICE coefficient loss function and an Intersection Over Union (IOU) metric. All code was developed with TensorFlow-2.1.0 on Python-3.6.7 within a Windows x64 environment on a GTX 2080 Ti GPU.

RESULTS: The mean age of participants was 63.2 ± 10.2 yrs, BMI: 29.67 ± 4.78 kg/m^2, 6% had KL= 0 or 1, 30% KL=2, 49% KL=3, and 15% KL=4. After more than 150 hours of training, top performing models tested as having a 1-DICE coefficient score of 0.196 (benchmark: less than 0.30) and a mean IOU of 0.681 (benchmark: higher than 0.50). The network also managed to segment BMLs in close proximity to one another at high spatial resolution with associated pixel confidence scores faithfully representing BMLs identified by manual segmentation.

CONCLUSION: Our UNET model managed to yield accurate high-resolution segmentations of discrete BMLs ranging in shapes with high fidelity. Future investigation of this model will focus on its performance in multi-class segmentation tasks, e.g. distinguishing BMLs from other features of interest, namely cysts.

SPONSOR: CIHR Project Grant PJT-156274
CORRESPONDENCE ADDRESS: andy.wong@uhnresearch.ca

A SCOPING REVIEW OF CONTEMPORARY METHODS OF GRADING RADIOGRAPHIC PATELLOFEMORAL OSTEOARTHRITIS

*Hill J.R., **Oei E.H.G., ***Crossley K.M., ***Menz H.B., **Macri E.M., ****Smith M.D., *****Wyndow N., ******Maclachlan L.R., ****Ross, M.H., *****Collins N.J.

*The University of Queensland, Brisbane, Australia & Ochsner Clinical School, New Orleans, LA, USA
**Erasmus MC University Medical Center, Rotterdam, The Netherlands
***La Trobe University, Melbourne, Australia
****The University of Queensland, Brisbane, Australia
*****The University of Queensland, Brisbane, Australia & La Trobe University, Melbourne, Australia
******Royal Brisbane and Women's Hospital, Brisbane, Australia & The University of Queensland, Brisbane, Australia

INTRODUCTION: Radiographic patellofemoral (PF) osteoarthritis (OA) affects 25% of the general population and is present in 39% of people with knee pain. Although radiographs are considered the gold standard in diagnosing PFOA, the majority of knee OA research has focused on the tibiofemoral joint (TFJ), resulting in a lack of development and universal adoption of a radiographic PFOA classification system. Systems designed to grade the TFJ are regularly applied to the patellofemoral joint (PFJ), although there has been no validation of their ability to accurately assess PFOA.

OBJECTIVE: Conduct a scoping review of the contemporary literature to identify commonly reported radiographic classification systems used to grade PFOA, and describe any reported measurement properties (e.g., reliability, validity) of these systems. We also sought to characterize which classification systems were most commonly used in different settings, disciplines, and geographical regions.

METHODS: In accordance with the PRISMA-ScR statement, we electronically searched eight common databases using keywords relating to "patellofemoral" and "radiograph". Two independent assessors screened each record for eligibility. English-language studies published between 2014-2018 were included if they reported on acquiring radiographs of the PFJ, described the method of radiograph acquisition, and reported radiographic grading of PFOA. Exclusion criteria included studies on subjects with a mean age less than ten years, non-human studies, cadaveric studies, and single subject studies. Quality appraisal was performed by two independent assessors using the COSMIN Risk of Bias Tool to Assess the Quality of Studies on Reliability and Measurement Error. Descriptive statistical analysis was performed.

RESULTS: Of the 18,678 records generated by the search, 177 articles met the inclusion criteria and data were extracted. Twenty-five classification systems to grade radiographic PFOA were reported. Of those specified, Kellgren-Lawrence (KL) (n=63, 37.3%), Iwano (n=24, 14.2%), and OARSI (n=16, 9.5%) were most prevalent. Of those articles specifying which radiographic projections were used to grade PFOA, axial projections (n=83, 82.2%) were most common, followed by lateral (n=41, 40.6%) and AP/PA (n=21, 20.8%) projections. Studies conducted in research settings (n=99, 55.9%) tended to use KL (n=32, 34.4%) and OARSI (n=15, 16.1%) criteria, whereas KL (n=25, 37.3%) and Iwano (n=15, 22.4%) were more frequently used in clinical settings (n=78, 44.1%). Although KL was most frequent across disciplines and regions, orthopaedics was the only specialty to report using the IKDC scale (n=13, 11.6%).

CONCLUSION: We found that multiple radiographic OA classifications systems are being used to grade PFOA, although few are specific to the PFJ. Progression of PFOA may not follow a similar disease course as OA in other joints, which may affect how generalizable non-PFJ specific grading systems are to the PFJ. Additionally, the variability in PFJ radiograph acquisition methods used may further influence the application of OA classification systems. For these reasons, we recommend that a reliable and valid PFOA radiographic grading system be developed using standardized PFJ radiograph acquisition techniques.

SPONSOR: none
DISCLOSURE STATEMENT: none
ACKNOWLEDGEMENT: none
CORRESPONDENCE ADDRESS: n.collins1@uq.edu.au

A LATENT MEASURE FOR SUBCHONDRAL BONE MINERAL DENSITY

*Ha E., *Wong A.K.

*Dalla Lana School of Public Health, University of Toronto, Toronto, Canada
*Joint Department of Medical Imaging, University Health Network, Toronto, Canada

INTRODUCTION: While computed tomography (CT) can provide a detailed measure of subchondral bone mineral density (BMD), it is not part of the standard of care for knee OA. Structural properties could be estimated from magnetic resonance images, but they are unable to yield BMD measures. Potential correlates of subchondral BMD often collected as part of the standard of care could be used to represent the conceptual unity of subchondral BMD. Potential correlates of subchondral BMD may include hip BMD, previous fractures, previous falls, performance-based tests (e.g., 20m walk test, 30-second chair stand test), and anthropometrics. Using these known correlates of subchondral BMD, the objective of this study is to develop and validate a latent concept of subchondral BMD for scenarios when CT imaging is not available.

OBJECTIVES: 1) Develop a subchondral BMD latent construct based on its potential correlates, and 2) Validate the subchondral BMD latent construct against directly measured subchondral BMD.

METHODS: Data from the Bone Ancillary Study (48-month visit) of the OAI were used. In total, 1,068 men and women were included in this study (mean age: 65.2±9.2 years, 24.5% non-overweight; 3.71% previous fracture). Information on demographics and clinical characteristics were collected at the 48-month study visit using interviewer-administered questionnaires. DXA systems (Lunar Prodigy Advance, GE Lunar Corp., Madison WI, USA) with investigational knee (enCORE 2007 Version 11.20.068) or spine analysis software were deployed at each OAI clinical site. DXA-based tibial plateau BMD measurements were determined using a customized ROI. The height of the medial and lateral tibial ROIs was fixed at 20 mm while the ROI width in the medio-lateral direction was set as ½ the distance between the medial and lateral bone edges of the tibial plateau. Generalized linear models were used to identify significant and clinically relevant indicators of measured subchondral BMD. Confirmatory factor analysis (CFA) was used to develop a reflective latent construct for subchondral BMD using the potential correlates as indicator variables. Model fit was assessed using absolute (Standardized Root Mean Square Residual <0.08), parsimonious (Root Mean Square Error of Approximation <0.08), and incremental (Comparative Fit Index ≥0.90, Tucker Lewis Index ≥0.95) fit indices. Good model fit is an indication that the generated model is theoretically consistent with the data used. Structural equation modeling was used to measure the correlation between the latent construct of subchondral BMD and the directly measured subchondral BMD from DXA for each of the medial and lateral compartments to assess validity.

RESULTS: Subchondral BMD was significantly explained by the following indicator variables: femoral neck BMD, prior fractures, prior falls, 20m walk test pace, 30-second chair stand, height, and weight ($p<0.05$). Using these indicator variables, a subchondral BMD latent construct was developed. In CFA models of latent subchondral BMD, good model fit was observed and assessed by absolute (<0.08), parsimonious (<0.08) and incremental fit indices (≥0.95). The latent construct of subchondral bone was significantly correlated with each of directly measured subchondral BMD in the medial ($B=0.940, SE=0.024, p<0.0001$) and lateral ($B=0.885, SE=0.027, p<0.0001$) compartment by DXA.

CONCLUSION: Subchondral BMD can be represented as a reflective latent construct comprised of indicator variables commonly collected as part of the standard of care for musculoskeletal diseases, particularly in older adults recommended for osteoporosis screening. This study provides a deeper understanding of correlates of subchondral bone and offers a unique and accessible method to estimate subchondral BMD status for those with suspected knee osteoarthritis.

SPONSOR: Canadian Institutes for Health Research (CIHR) PJT166012 and PJT156274
DICLOSURE STATEMENT: EH and AW have no conflicts of interests to declare.
ACKNOWLEDGMENT: The Osteoarthritis Initiative (OAI) is a public-private partnership comprised of five contracts (N01-AR-2-2258; N01-AR-2-2259; N01-AR-2-2260; N01-AR-2-2261; N01-AR-2 2262) funded by the National Institutes of Health, a branch of the Department of Health and Human Services, and conducted by the OAI Study Investigators.
CORRESPONDENCE ADDRESS: andy.wong@uhnresearch.ca

CLINICAL APPLICATION OF ULTRA-HIGH FIELD MRI FOR WRIST IMAGING: A MULTI-READER COMPARISON OF 3T AND 7T

*Heiss R., **Weber M.A., ***Schmitt R., ****Rehnitz C., *****Laqmani A., ******Sternberg A., *******Ellermann J., *Nagel A.M., *Arkudas A., *******Horch R., *Uder M., *Roemer F.W.

*Universitätsklinikum Erlangen & Friedrich-Alexander University Erlangen-Nürnberg, Erlangen; Germany
**Institute of Diagnostic and Interventional Radiology, University Medical Center Rostock, Rostock, Germany
***LMU München - Campus Großhadern, Munich, Germany
****Universitätsklinikum Heidelberg, Heidelberg, Germany
*****Universitätsklinikum Hamburg-Eppendorf, Hamburg, Germany
******Medizinisches Versorgungszentrum am Rotes Kreuz Krankenhaus, Bremen, Bremen, Germany
*******University of Minnesota Medical School, Minneapolis, MN, USA

INTRODUCTION:
Potential advantages of ultrahigh-field MRI in musculoskeletal imaging include the visualization of small anatomic structures due to high spatial resolution. The wrist is one of the anatomic regions where these advantages may be clinically relevant. A direct comparison of 3T vs 7 T MRI for visualization of multi-tissue assessment of the wrist joint has not been performed to date.

OBJECTIVE: To compare 7T and 3T MRI of the wrist based on semiquantitative scoring of multiple joint tissues of seven readers and to assess overall image quality.

METHODS: Twenty-five subjects and 25 patients with chronic wrist pain were examined at 3T and 7T on the same day (T1 turbospinecho coronal, fat saturated PD turbospinecho coronal, T2 turbospinecho transversal, Table 1). Image data were evaluated separately for 3T and 7T via online scoring templates by a seven-member MSK radiologist expert panel. Image quality, presence of artifacts, homogeneity of fat saturation, visualization of cartilage, TFCC, SL and LT ligament structures were semiquantitatively assessed using Likert scales. Pathologies in subjects and patients were recorded using the 7T images, and interreader reliability was determined.

	3T		7T	
	voxel size	TA	voxel size	TA
T1 tse cor	0.2×0.2×2.0	04:29	0.1×0.1×2.0	03:16
PD tse fs cor	0.2×0.2×2.0	04:02	0.2×0.2×2.0	04:02
T2 tse tra	0.2×0.2×2.0	04:15	0.2×0.2×2.0	04:15

Table 1. Acquisition parameters

RESULTS: Overall image quality of PD cor fs was found to be significantly superior on 7T compared to 3T (difference mean (Dm) -0.16, p=0.002), with T1 cor (Dm 0.42, p<0.001) and T2 tra (Dm 0.40, p<0.001) being superior for 3T. 3T MRI was superior regarding presence of artifacts for all sequences (p = 0.001) (Fig.1) and for the quality of fat saturation (p < 0.001). Visualization of cartilage was significantly better at 7T (Dm -0.30, p < 0.001) (Fig. 2). Visualization of TFCC, SL and LT band was significantly superior for 3T (Dm 0.24, p<0.001; Dm 0.06, p=0.048, Dm 0.046, p=0.042, respectively). The reliability of all 7 readers showed moderate to substantial agreement for the detected pathologies (κ = 0.21-0.64).

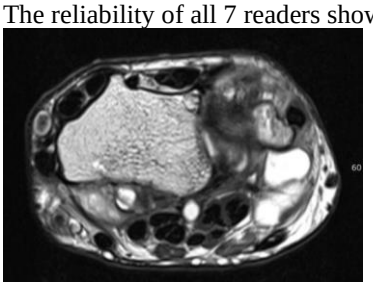


Fig. 1. Increased field inhomogeneities at 7T (right image)

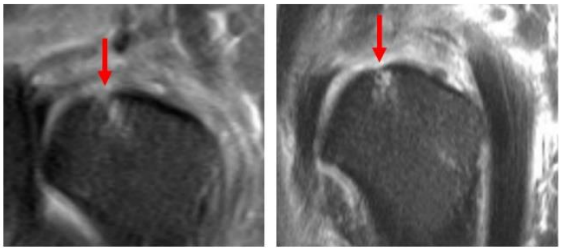


Fig 2. Superior visualization of focal cartilage lesion at 7T (right image)

CONCLUSION:
Depending on the sequence and anatomic structure, ultrahigh-field imaging at 7T at wrist has potential advantages over 3T MRI, particularly with regard to cartilage imaging. 3T MRI was superior to 7T in terms of artifacts and fat saturation.

SPONSOR: Supported by research grant of the German Roentgen Society (Deutsche Röntgengesellschaft–DRG).
DICLOSURE STATEMENT: Dr. Roemer is shareholder of BICL, LLC. Dr. Roemer is consultant to Calibr - California Institute for Biomedical Research. None of the other co-authors declared potential competing interests.
CORRESPONDENCE ADDRESS: rafael.heiss@uk-erlangen.de

RELIABILITY OF AUTOMATED KNEE JOINT SPACE WIDTH MEASURES USING IB LAB KOALA: DATA FROM THE OSTEOARTHRITIS INITIATIVE

*Paixao, T., *Salzlechner C., *Hummer A., *DiFranco M.

*ImageBiopsy Lab, Vienna, Austria

INTRODUCTION: Femorotibial JSN is a frequently used indicator to assess OA in the knee. Tracking JSN progression in a patient over time can be challenging due to qualitative reproducibility of knee radiographs. In addition, assessment of severe bone-on-bone OA can be difficult due to the small space available.

OBJECTIVE: Assess the reliability of fully automated longitudinal JSW measurements using IB Lab KOALA with OAI ground truth and stratifying by JSN. In addition, assess the reliability of KOALA JSW measures when considering different beam angles as reported in OAI, and compare KOALA JSW measures to OAI on patients with reported bone-on-bone contact in the medial compartment.

METHODS: A testing set of 1149 patients unseen by the KOALA algorithm was used for all assessments. Linear regressions of longitudinal minimum JSW (minJSW) measures over time were performed independently for OAI and KOALA on individuals with at least three timepoints and paired differences of the slopes between KOALA and OAI analyzed. For each knee for which a 12-month follow-up was available (knees at 48, 72 and 96 months were excluded), JSN and its annual change (ΔJSN) was compared with the annual change in minJSW for OAI and KOALA.
Additionally, bootstrapped regressions of paired medial minJSW measurements of individual knees from KOALA and OAI were stratified by binned OAI-reported beam angle into Low (< 8°, n=571), Medium (8° to 12°, n=5280) and High (> 12°, n=1183) bins.
Finally, fixed JSW measures from OAI and KOALA were compared on 87 knees with OAI-reported medial bone on bone and minJSW=0. We directly compare the OAI fixed width JSW measures on the medial side (JSW150 - JSW300) with the four JSW measures from KOALA.

RESULTS: Longitudinal minJSW change between KOALA and OAI had a mean difference of -0.0019 +/- 0.0079 mm [95% CI: -0.0156 mm; 0.0120 mm] medial and -0.0004 +/- 0.0109 mm [95% CI: -0.0155 mm; 0.0211 mm] lateral. For annual change of JSW stratified by JSN and ΔJSN, no significant differences were found for any JSN or ΔJSN.
For regression of medial minJSW between KOALA and OAI, slopes were 1.05 +/- 0.03 [95% CI: 1.00; 1.10] for Low beam angles, 1.06 +/- 0.06 [95% CI: 0.98; 1.19] for Middle beam angles, and 1.04 +/- 0.05 [95% CI: 0.97; 1.16] for High beam angles.
In the 87 knees with medial bone on bone in OAI and minJSW of 0, KOALA reported minJSW=0 on 66%, minJSW <= 0.25 mm on 72%, minJSW <= 0.5 mm on 83%, minJSW <= 0.75 mm on 93% and minJSW <= 1 mm on 97%.

CONCLUSION: We show that automated JSW measures from KOALA are reliable as compared to OAI ground truth in a longitudinal setting and when stratified by JSN and beam angle bins. KOALA can detect bone on bone medial compartment while sometimes reporting marginal joint space where OAI ground truth does not.

SPONSOR: IB Lab GmbH, Vienna, Austria
DISCLOSURE STATEMENT: CS, AH and MD are employees of IB Lab GmbH. TP was employed by IB Lab GmbH when carrying out the work on this study.
CORRESPONDENCE ADDRESS: m.difranco@imagebiopsy.com

CONTEMPORARY METHODS OF ACQUIRING PATELLOFEMORAL JOINT RADIOGRAPHS: A SCOPING REVIEW

*Hill J.R., **Oei E.H.G., ***Crossley K.M., ***Menz H.B., **Macri E.M., ****Smith M.D., *****Wyndow N., ******Maclachlan L.R., ****Ross, M.H., *****Collins N.J.

*The University of Queensland, Brisbane, Australia & Ochsner Clinical School, New Orleans, LA, USA
**Erasmus MC University Medical Center, Rotterdam, The Netherlands
***La Trobe University, Melbourne, Australia
****The University of Queensland, Brisbane, Australia
*****The University of Queensland, Brisbane, Australia & La Trobe University, Melbourne, Australia
******Royal Brisbane and Women's Hospital, Brisbane, Australia & The University of Queensland, Brisbane, Australia

INTRODUCTION: Patellofemoral joint (PFJ) pathology is prevalent and associated with significant morbidity. Although protocols for PFJ radiograph acquisition exist, standardized methods to acquire radiographs of the PFJ have not been universally adopted, resulting in variances in patient positioning, weight-bearing status, knee flexion angle, and radiograph beam direction, leading to inconsistent joint assessment.

OBJECTIVE: Conduct a scoping review of the literature to determine commonly described methods of acquiring PFJ radiographs, and evaluate the methods used in specific settings, disciplines, regions, and for different pathologies.

METHODS: Adhering to PRISMA-ScR guidelines, we searched common databases using keywords related to the concepts of 'patellofemoral' and 'radiograph'. Two independent assessors screened each record for eligibility. Studies published between 2014-2018 describing radiographic acquisition methods of the PFJ were included. Studies of non-human subjects, pediatric populations under 10 years, cadaver studies, and single-subject studies were excluded. Qualitative synthesis of radiographic methods was performed.

RESULTS: From 15,053 unique records, 549 articles were included. Four radiographic projections were reported (lateral (84.3% of articles), axial (68.5%), anteroposterior/posteroanterior (66.3%), and lower-limb alignment (21.1%)). We found marked variability for each projection in relation to views acquired (e.g., there were 13 different views described for axial projections) and techniques used (e.g., patient positioning, radiograph beam direction). Acquisition techniques varied across disciplines, settings, regions, and pathologies without clear patterns emerging.

CONCLUSION: Our findings highlight considerable variability in PFJ radiograph acquisition methods in contemporary literature. The lack of standardization of PFJ radiograph acquisition may make it difficult to accurately assess and estimate true population prevalence of PFJ pathology.

SPONSOR: none
DISCLOSURE STATEMENT: none
ACKNOWLEDGEMENT: none
CORRESPONDENCE ADDRESS: n.collins1@uq.edu.au

EVALUATION OF FULLY AUTOMATED KNEE ALIGNMENT ASSESSMENT IN LOWER EXTREMITY RADIOGRAPHS

*,**Simon S., **,***Schwarz G, *,**Aichmair A., *Frank B.J.H., ****Hummer A., ****DiFranco M.D., **,*****Dominikus M., *,**Hofstätter J.

*Michael Ogon Laboratory for Orthopaedic Research, Orthopaedic Hospital Vienna-Speising, Austria
**Department of Orthopaedic Surgery, Orthopaedic Hospital Vienna-Speising, Austria
***Department of Orthopedics and Trauma Surgery, Medical University Vienna, Austria
****ImageBiopsy Lab, Vienna, Austria
*****Sigmund Freud University Vienna, Austria

INTRODUCTION: Long-leg anteroposterior radiographs (LLR) are acquired as the standard method to assess knee alignment and limb length discrepancies for surgical planning and post-operative assessment. Currently, radiologists and surgeons perform these measurements using manual annotations or computer-assisted tools. There is no standardization in technique of measurements or in the use of different software applications, which leads to high intra- and inter-reader variability, a lack of reproducibility and length measurement calibration errors.

OBJECTIVE: Fully-automated measurements (FAM) could help deliver standardized measurements and calibration. The main purpose of this study is to evaluate the feasibility and reproducibility of a fully automated artificial intelligence (AI) software to measure seven angles and four lengths from long-leg anteroposterior radiographs (LLR).

METHODS: We performed a single center analysis on 295 LLRs. A FAM and two clinicians measured all images with repeats after a washout period of at least 4 weeks. All measurements were compared and interchangeability, mean absolute deviation (sMAD) and intraclass correlation (ICC) were calculated. Moreover, clinicians self-reported grade of difficulty for each image.

RESULTS: FAM vs. mean reader revealed an sMAD between 0.39 to 2.19° for all angles and 1.45-5.00 mm for all lengths. The FAM showed a good reliability in all lengths and angles (ICC≥0.87) compared to the mean of readers. Interchangeability assessment comparing the FAM to the mean reader revealed an equivalence index (γ) of 0.54° for the hip-knee-ankle (HKA) angle, 0.70 mm for uncalibrated leg length, and 6.70 mm for calibrated leg length. We found excellent inter- and intra-reader (ICC≥0.9) correlation in all measurements. For the FAM there was no intra-FAM variability in angles and in lengths there was an ICC>0.99. FAM is three times faster than clinicians at a high level of reliability.

CONCLUSION: FAM of LLR yielded accurate results. The potential to automate tasks by AI able to produce standardized outputs in a big amount of data in a short time.

SPONSOR: IB Lab GmbH, Vienna, Austria
DISCLOSURE STATEMENT: AH and MD are employees of IB Lab GmbH. Michael Ogon Laboratory for Orthopaedic Research, Orthopaedic Hospital Vienna-Speising receives a research grant from IB Lab GmbH.
CORRESPONDENCE ADDRESS: m.difranco@imagebiopsy.com

ASSOCIATION OF MENISCAL ROOT TEARS WITH KNEE OA: INITIAL FINDINGS FROM THE OAI COHORT

*Bharadwaj U.U., *Ramezanpour S., ****Schiro S., **Feeley B., **Lansdown D., ***Lynch J., *Joseph G.B., *Link T.M.

*Department of Radiology and Biomedical Imaging, University of California, San Francisco, USA
**Department of Orthopaedic Surgery, University of California, San Francisco, USA
***Department of Epidemiology and Biostatistics, University of California, San Francisco, USA
****Department of Medicine and Surgery (DiMeC), Section of Radiology, University of Parma, Italy

INTRODUCTION: Meniscal root tears significantly impact knee joint health, and account for 10-21% of all types of meniscal tears. MRI studies have previously shown associations between root tears and cartilage damage; however, prior studies also report presence of meniscal root tears in patients with mild degenerative changes of the knee, implying the need for more detailed studies to identify morphological characteristics of meniscal root tears associated with knee OA.

OBJECTIVE: To evaluate the association of meniscal root tear morphology, namely partial and complete root tears, with knee OA using WORMS to quantify OA outcomes.

METHODS: Participants (n=48: 35 female, 13 male) from the OAI cohort with root tears (age: 63.05 +/- 1.4 years; BMI: 29.84 +/- 0.85 kg/m^2) and controls (age: 63.6 +/- 2.3 years; BMI: 31.1 +/- 1.4 kg/m^2) matched for age, BMI were selected at a single timepoint. MRIs of these patients were evaluated to identify partial or complete tears of the anterior/posterior roots of the medial and lateral meniscus, meniscal extrusion, ACL tears, and osteophytes. Severity of OA was characterized using WORMS grading for cartilage, BM lesions, and meniscus. Linear regression models (adjusted for age, sex, and BMI (at baseline) were used to identify associations between root tear morphology (partial or complete tear) and WORMS.

RESULTS: Differences in age, sex, and BMI between patients with meniscal root tears (n=35) and those without any noted tears (n=13) were not statistically significant ($p > 0.1$). The assessed tears included both partial tears (n=24) as well as complete tears (n=11). The medial meniscus had 17 tears (n=4 complete and n=13 partial) and the lateral meniscus had 18 tears (n=7 complete and n=11 partial). Linear regression indicated strong association between the type of root tear (partial or complete) and knee OA with statistically significant differences in WORMS grades between partial vs control and complete vs control as shown in Table 1.

Parameter	Controls		Partial		Complete		Partial vs Control	Complete vs Control
	mean	SE	mean	SE	mean	SE	p-value	p-value
WORMS_MFC_Cartilage	0.84	0.38	1.49	0.29	1.75	0.42	0.0001	0.0001
WORMS_MT_Cartilage	0.32	0.41	1.21	0.31	1.66	0.45	0.0001	0.0001
WORMS_MedialMeniscus	0.14	0.47	1.49	0.33	2.05	0.50	0.0001	0.0001
WORMS_LFC_Cartilage	0.63	0.24	0.53	0.18	0.42	0.26	0.01	0.12
WORMS_LT_Cartilage	0.79	0.32	0.92	0.24	1.46	0.36	0.001	0.0001
WORMS_LateralMeniscus	0.66	0.58	0.73	0.41	2.04	0.62	0.082	0.002

Table 1: Adjusted mean +/- standard error (SE) of WORMS grades across types of root tears.

CONCLUSION: This preliminary study suggests that the type of meniscal root tear is associated with severity of knee OA measured with WORMS. With the notable exception of WORMS_LFC_Cartilage, complete root tears had higher WORMS scores than the control group (statistically significant) as well as partial tears. A detailed study characterizing the association of root tear morphology with knee OA may provide additional insight for subsequent management. Since partial tears may eventually progress into complete tears, a longitudinal study assessing root tear morphology and progression of knee OA would be valuable.

SPONSOR: Funded by NIH/NIAMS grants R01AR064771 and R01AR078917
DICLOSURE STATEMENT: None
CORRESPONDENCE ADDRESS: Upasana.Bharadwaj@ucsf.edu

3D MRI ANALYSIS FOR CARTILAGE IN ANTERIOR CRUCIATE LIGAMENT DEFICIENT KNEES USING RADIALLY PROJECTED IMAGES

*Ozeki N., **Suzuki K., **Masumoto J., *Sekiya I.

*Center for Stem Cell and Regenerative Medicine, Tokyo Medical and Dental University, Tokyo, Japan
**Fujifilm Corporation, Tokyo, Japan

INTRODUCTION: Cartilage injury often occurs after anterior cruciate ligament (ACL) injury, and it has been reported that knee instability or duration from injury to surgery affects the incidence of the injury. We have been developing a new 3D MRI analysis system that enables to reconstruct 3D images of the cartilage in the knee joint semi-automatically, and it is useful to quantitate the condition of cartilage by projected cartilage area ratio with intended thickness (Hyodo, et al. J Bone Joint Surg OA 2019). Then we have improved the 3D-reconstructed femoral cartilage to project radially onto the 2D plane automatically (Aoki, et al. In submission).

OBJECTIVE: To examine the radially projected cartilage area ratio in ACL injured knees and to investigate the factors that were correlated with cartilage injury.

METHODS: This study included 61 patients who had 3T MRI (Achieva 3.0TX, Philips, pixel spacing: 0.31mm×0.31mm, slice thickness: 0.6mm) before ACL reconstruction. Average age was 25±10 years old (males/female; 30/31). After 3D reconstruction, femoral cartilage was radially projected and tibial platearu was projected to a flat surface, and each ROI was set to 9 segments (Figure 1, 2). In this study, the projected cartilage area ratio with cartilage thickness of 1.5 mm and over was evaluated. Correlations between the cartilage area ratio and the following factors were analyzed using multiple linear regression analysis ; age, duration from injury to surgery (<90 days), the incidence of giving way (<5 times), anterior instability measured by knee arthrometer (mm) , pivot shift test grade quantified at 0 to 7 (<4), Lysholm score, and Tegner Activity Scale.

RESULTS: Several factors had significant effects on certain projected cartilage ratio, duration from injury to the surgery; miLF (p=0.014), mcLF(p=0.037), meLF(p=0.012), pcMT (p=0.032), incidence of giving way; aiLF (p=0.026), acLF(p=0.036), meLF(p=0.048), mdTrF (p=0.048), peLT(p=0.009), Lysholom score; meMF (p=0.046), Tenger Activity Scale; miTrF (p=0.028).

CONCLUSION: Radially projected cartilage area ratio revealed the factors that had effects on the cartilage status after ACL injury. Duration from injury to surgery and the incidence of giving way significantly affected on the multiple cartilage areas especially in the lateral compartment of the knee joint.

SPONSOR: Ichiro Sekiya (Japan Agency for Medical Research and Development [AMED])
DISCLOSURE STATEMENT: None
ACKNOWLEDGMENT: None
CORRESPONDENCE ADDRESS: ozeki.arm@tmd.ac.jp

Figure 1

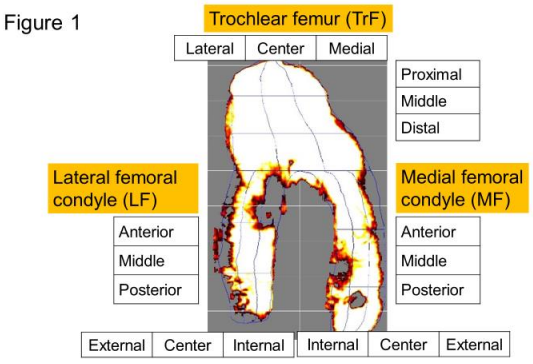
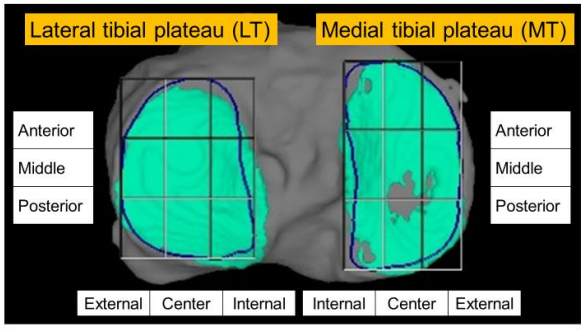


Figure 2



PROGNOSTIC MODEL IN EARLY OSTEOARTHTISI BASED ON MRI-BASED STRUCTURAL PHENOTYPES

*Lee J.J., *Nimiri N.K., *Astuto B., *Link T., *Majumdar S., *Pedoia V.

*Department of Radiology and Biomedical Imaging, University of California San Francisco

INTRODUCTION: The availability of a personalized model to predict the risk of structural OA incidence would be useful by aiding the selection of participants for clinical trials of potential disease modifying molecules. Also, such a model could be used to assess effectiveness of treatments compared to natural disease progression ultimately leading to personalized medical management. Recently MRI-based scoring system has been proposed to quickly assess eligibility for clinical trial inclusion. Building upon our previous study that utilized deep learning to evaluate the phenotypes in large scale, this study developed a model to predict the risk of the development of structural OA.

OBJECTIVE: We investigated 1) whether the risk of structural OA development varies by proposed phenotypes and 2) how interactions of among phenotypes, in case of multiple phenotype presence, alters the risk of structural OA development.

METHODS: The study population was the OAI participants who did not have radiographic OA at baseline visit (N = 2,328). We excluded knees that had definite tibial-femoral osteophytes (OARSI atlas grade 1-3) and pain, aching or stiffness on most days of a month in past year before baseline visit. If a participant did not have radiographic OA on both knees, then we chose the dominant knee which a participant used to kick a ball in isometric strength test at enrollment visit. Using the previously developed deep learning model, we automatically assessed morphological phenotypes, based on definitions from the Rapid OsteoArthritis MRI Eligibility Score of subchondral bone, meniscus/cartilage, inflammatory, and hypertrophy phenotypes.
The Cox proportional hazard model was used to derive a model based on structural phenotypes. At 12-, 24-, 36-, 48-, 72-, and 96- month follow-up visits, MRIs were acquired, thus phenotype predictions were available. Therefore, we examined structural phenotype as time-dependent variable in all analyses. Development of radiographic OA was treated as a failure for the purpose of survival analysis. Specifically, radiographic OA was defined as Kellgren-Lawrence (KL) scale greater or equal to 2 (presence of definite osteophytes and joint space narrowing). The initial time point for survival modeling was the date the first knee MRIs were acquired. Participant contributed person-time until they underwent knee replacement surgery, withdrew from the study, or reached the end of the study, whichever occurred first. A total of 12 variables (described in Table 1) were examined individually then collectively. In selecting variables, we used exhaustive model search using Akaike information criteria (AIC). The selected model was validated with holdout subset of data. After validation, the final model was fitted to the combined dataset.

RESULTS: Univariate survival modeling result is shown in Table 1 with descriptive statistics. All of four presence of phenotypes were predictive of risk of structural OA. From 12 variables, the exhaustive search found BMI and knee injury history were redundant. The remaining 10 variables were examined further by assessing two-factor interactions, and included if it were found to be statistically significant at the 0.05 level. Table 2 shows the estimated hazard ratios for the multivariate Cox model. When adjusted for age at enrollment, sex, WOMAC pain score of both knees (time-dependent), presence of Heberden nodes, and knee alignment, the hazard ratios (95% CI) of structural OA for inflammatory-, meniscus/cartilage-, and subchondral bone- phenotype are 3.79 (2.97 – 4.83), 2.06 (1.48- 2.88), and 1.99 (1.74-2.28), respectively (Figure 1). The effect of hypertrophy phenotype on risk structural OA were estimated to be negative, but it did not reach the level of statistical significance. The effect of hypertrophy phenotype was positive for female, but the confidence interval was too wide to reach the conclusion. The hazard ratio of presence of both meniscus/cartilage and subchondral bone phenotype is 2.48 compared to absence of such phenotypes, 1.24 times higher than the bone phenotype alone, 1.21 times higher than the meniscus/cartilage phenotype alone. The hazard ratio of presence of meniscus/cartilage and inflammatory phenotype is 2.41 relative to absence of such phenotypes, 1.17 relative to the meniscus/cartilage phenotype, and 0.64 relative to inflammatory phenotype alone. Final model can be used for personalized structural OA risk prediction. An example is shown in Figure 2.

CONCLUSION: Four phenotypes are predictive of structural OA. We found interaction between subchondral bone and meniscus/cartilage and interaction between inflammatory and meniscus/cartilage.

SPONSOR: NIH, Grant No 0000.
DICLOSURE STATEMENT: Nothing to disclose.
CORRESPONDENCE ADDRESS: jennylee.stat@gmail.com

ANATOMICAL PREVALENCE OF CARTILAGE DAMAGE IN KNEES WITH RADIOGRAPHIC DISEASE SEVERITY KL 2 AND 3: A SUBREGIONAL ANALYSIS FROM THE MOST STUDY

*,**Roemer F.W., *Felson D.T., ***Stefanik J.J., *Rabasa G., *Wang N., *Crema M.D., *Neogi T., ****Nevitt M.C., *****Torner J., ******Lewis C.E., *Guermazi A.

*Boston University, Boston, MA, USA
**Friedrich-Alexander University Erlangen-Nürnberg (FAU) and Universitätsklinikum Erlangen, Erlangen, Germany
***Northeastern University, Boston, MA, USA, ****University of California, San Francisco, CA, USA
*****University of Iowa, Iowa City, IA, USA, ******University of Alabama at Birmingham, Birmingham, AL, USA

INTRODUCTION: Imaging plays an important role in defining structural disease severity and potential suitability of patients recruited to DMOAD trials. To be eligible for inclusion into a DMOAD trial from a structural perspective, knees with KL grades 2 and 3 are typically included. These knees exhibit definite structural disease, are at risk for progression, but are not considered end-stage. To date anatomical distribution of cartilage damage based on MRI and ordinal grading approaches has not been well described. In addition, BMLs are one of the recognized pain-generating features of OA and subregions with both BMLs and cartilage damage have increased risk of structural progression. Thus, knees or compartments exhibiting both cartilage damage and adjacent BMLs may be particularly relevant for targeted treatment approaches.

OBJECTIVE: To assess the prevalence 1) of any cartilage damage in 14 articular subregions and 2) of any cartilage damage and adjacent BMLs in the same articular subregion.

METHODS: The Multicenter Osteoarthritis (MOST) study is a prospective cohort study of individuals 60-79 years with or at risk for knee OA. All baseline MRIs of subjects included in the MOST study with radiographic disease severity KL 2 and 3 were included. MRI was performed using a 1.0 T extremity system at the two clinical sites using axial and sagittal proton-density weighted sequences and a coronal STIR sequence. Knee MRIs were read by two experienced MSK radiologists for cartilage damage from 0 to 6 regarding subregional area and full thickness involvement according to the WORMS scoring system in 14 articular subregions: (4 subregions of the patellofemoral joint (PFJ): medial patella (mP), lateral patella (lP), anterior medial femur (amF), anterior lateral femur (alF); 5 subregions of the medial TFJ: central medial femur (cmF), posterior medial femur (pmF), anterior medial tibia (amT), cental medial tibia (cmT) and posterior medial tibia (pmT); 5 subregions in the lateral lTFJ: central lateral femur (clF), posterior lateral femur (plF), anterior lateral tibia (alT) central lateral tibia (clT), posterior lateral tibia (plT). BMLs were assessed according to WORMS from 0-3 in the same subregions. The ranks of affected subregions for KL2 and 3 knees combined were assessed for all 14 subregions as well as for those knees with any cartilage damage and concomitant BMLs in the same subregion.

RESULTS: 445 knees had a radiographic disease severity of KL2 and 317 knees were KL3 at the MOST baseline visit. Concerning KL2 and 3 knees combined the ranking of subregions in terms of prevalence of any cartilage damage in these knees was: 1.) cmF (79.7%) 2.) mP (72.8%) 3.) cmT (68.8%) 4.) lP (57.9%) and 5.) pmF (53.7%). For those subregions with adjacent BMLs the prevalence of the most affected subregions was: 1.) cmT (28.2%) 2.) cmF (26.6%) 3.) pmF (17.0%) 4.) alF (15.2%) and 5.) clT (12.1%). Details are presented in **Table 1**.

Any cartilage damage and any cartilage damage with adjacent BML, full thickness cartilage damage and full thickness cartilage damage with adjacent BML all subregions KL2 and 3 combined

		ANY Cartilage Damage		ANY Cartilage Damage AND BML		Full Thickness Cartilage Damage		Full Thickness Cartilage Damage AND BML	
Compartment	Subregion	N (%)	Rank	N (%)	Rank		Rank	N (%)	Rank
mTFJ	cmT	524 (68.7)	3	215 (28.2)	1	305 (40.0)	1	172 (22.5)	1
	cmF	607 (79.6)	1	203 (26.6)	2	302 (39.6)	2	151 (19,8)	2
	pmF	409 (53.7)	5	129 (17.0)	3	61 (8.0)	10	31 (4.1)	8
	pmT	159 (20.8)	13	23 (3.0)	13	24 (3.2)	13	8 (1.0)	13
	amT	190 (24.9)	11	51 (6.7)	8	45 (5.9	12	19 (2.5)	8
lTFJ	clT	284 (37.2)	9	92 (12.1)	5	106 (13.9)	6	63 (8.3)	4
	clF	340 (44.6)	7	53 (7.0)	7	97 (12.7	7	28 (3.7)	9
	plF	167 (21.9)	12	50 (6.6)	10	48 (6.3)	11	26 (3.4)	10
	plT	229 (30.1)	10	44 (5.8)	11	88 (11.6)	8	25 (3.3)	11
	alT	56 (7.3)	14	4 (0.5)	14	21 (2.8)	14	3 (0.3)	14
PFJ	mP	541 (72.8)	2	49 (6.6)	9	200 (26.9)	4	32 (4.2)	7
	amF	381 (50.1)	6	29 (3.8)	12	66 (8.8)	9	15 (2.0)	12
	lP	430 (57.9)	4	80 (10.8)	6	201 (27.1)	3	68 (8.9)	5
	alF	318 (41.8)	8	116 (15.2)	4	167 (21.9)	5	94 (12.3)	3

CONCLUSION:
In KL2 and 3 knees cartilage damage is most observed in the cmF (~80%) subregion followed by the mP and cmT (both ~70%) subregions. The most affected subregion with cartilage damage and concomitant BMLs is the cmT (30%) followed by the cmF (27%) and pmF (17%). The heterogeneity knees with structural disease severity KL2 and 3 show a wide spectrum of subregional structural involvement and subregional disease burden (e.g. concomitant presence of cartilage damage and BMLs).

SPONSOR: Supported by NIH grants from the National Institute of Aging to Drs. Lewis (U01-AG-18947), Torner (U01-AG-18832), Nevitt (U01-AG-19069), and Felson (U01-AG-18820).
DICLOSURE STATEMENT: AG is Consultant to Merck Serono, AstraZeneca, Pfizer, Novartis, Regeneron, and TissueGene and is shareholder of BICL, LLC. FWR is shareholder of BICL, LLC. Dr. Roemer is consultant to Calibr.
CORRESPONDENCE ADDRESS: frank.roemer@uk-erlangen.de

ANY CHANGE IN SEMI-QUANTITATIVE MOAKS CARTILAGE SCORES CORRESPONDS TO CARTILAGE THICKNESS LOSS: DATA FROM THE FNIH BIOMARKER CONSORTIUM

*Roemer F.W., **Maschek S., **Wisser A., ***Guermazi A., ****Hunter D.J., **Eckstein F., **Wirth W.

*Department of Radiology, University of Erlangen, Boston University & BICL, Boston, MA, USA
**Paracelsus Medical University, Salzburg, Austria & Chondrometrics GmbH, Freilassing, Germany
***Department of Radiology, Boston University & BICL, Boston, MA, USA
****Rheumatology Department, Royal North Shore Hospital and Institute of Bone and Joint Research, Kolling Institute, University of Sydney, Sydney, NSW, Australia

INTRODUCTION: Both semi-quantitative (SQ) MOAKS cartilage scores and quantitative (Q) cartilage thickness measurements are used as an outcome measure in observational studies and interventional trials. In order to increase sensitivity to change, so-called within-grade changes not fulfilling criteria of a full grade change have been introduced in SQ assessment. While the validity of these changes has been shown, the association between longitudinal change of SQ-defined within-grade and full grade changes and longitudinal Q cartilage loss has not been established.

OBJECTIVE: To study the association between full-grade and within-grade change in MOAKS cartilage scores and concurrent change in quantitative cartilage thickness measurements over 24 months.

METHODS: The 600 participants from the OAI FNIH study had MOAKS assessments and quantitative cartilage thickness measurements (age: 62y, BMI: 31kg/m², 59% female). MOAKS assesses cartilage in a two-digit fashion taking into account the area extent and the extent that is full thickness loss (possible grades: 0.0, 1.0, 1.1, 2.0, 2.1, 2.2, 3.0, 3.1, 3.2 and 3.3). The maximum MOAKS cartilage score (MOAKSmax), the maximum MOAKS area extent score (MOAKSext, 1st component of MOAKS cartilage scale), and the maximum MOAKS cartilage damage full thickness score (MOAKSft, 2nd MOAKS cartilage component) in each 3 tibial and 1 central femoral MOAKS subregion were determined at baseline and two-year follow-up for the medial and lateral compartment. Medial and lateral compartment cartilage thickness (MFTC/LFTC) change over the subsequent two years were stratified by ipsicompartmental change in MOAKSmax, MOAKSext, and MOAKSft score. Within-grade changes were assessed in knees without full-grade change in comparison to knees without any change. Between-group comparisons were performed using ANCOVA with adjustment for age, sex, and BMI. Results were presented as adjusted mean difference and 95% confidence intervals.

RESULTS: Knees with any full-grade increase in MOAKSmax in the medial compartment (n=186) showed more MFTC cartilage loss than knees that remained stable (n=412, -0.18mm, 95% CI: [-0.21, -0.14]mm). This difference was less pronounced in the lateral compartment (n=50 vs. n=547, -0.05mm, [-0.08, -0.02]mm). In the medial compartment, knees with an increase in medial MOAKSext by 1 grade (n=57, -0.16mm, [-0.21, -0.10]mm) and knees with an increase in medial MOAKSft by 1 (n=85, -0.15mm, [-0.20, -0.11]mm) or 2 grades (n=46, -0.17mm, [-0.23, -0.11]mm) showed greater MFTC loss than knees without such change (MOAKSext/MOAKSft: n=520/468). In the lateral compartment, LFTC cartilage thickness loss was elevated in knees with increase in MOAKSext but not in knees with increase in MOAKSft (data not shown).
Knees with isolated within-grade changes in the medial compartment (n=64) showed a greater loss in MFTC cartilage thickness compared to those without any change (n=412, -0.08mm, [-0.11, -0.04]mm), while knees with isolated lateral within-grade changes (n=13) did not.

CONCLUSION: Both full-grade and within-grade changes in MOAKS cartilage scores corresponded with ipsicompartmental cartilage thickness loss. In the medial compartment, concurrent change in cartilage thickness was comparable for change in the MOAKSext and MAOKSft change dimensions. The effect was less pronounced for MOAKS within-grade changes but still significant, supporting the clinical validity of within-grade assessment and its relevance in a longitudinal study design.

DISCLOSURE STATEMENT: WW: Consulting for Galapagos; DJH: consulting for Pfizer, Lilly, TLCBio, Novartis, Tissuegene, Biobone; FE: consulting for Merck KGaA, Abbvie, Samumed, Kolon-Tissuegene, Servier, Galapagos, Roche, Novartis, ICM and HealthLink. AG: consulting for Pfizer, Kolon TissueGene, Novartis, AstraZeneca, Merck Serono and Regeneron.
ACKNOWLEDGMENT: Foundation for the NIH (FNIH), OAI participants, staff, and investigators.
CORRESPONDENCE ADDRESS: Frank.Roemer@uk-erlangen.de

WBCT PROVIDES AN INCREASED RATE OF DETECTION OF MENISCAL EXTRUSION COMPARED WITH MRI IN ADULTS WITH OR AT INCREASED RISK FOR KNEE OA

*,**Segal N.A., ***Roemer F.W., ****Lynch J.A., **Anderson D.D., *He J., ****Nevitt M.C., ***Guermazi A. for the MOST Study Group

*University of Kansas, Kansas City, KS, USA
**The University of Iowa, Iowa City, IA, USA
***Boston University School of Medicine, Boston, MA, USA
****University of California-San Francisco, San Francisco, CA, USA

INTRODUCTION: Meniscal extrusion may be missed on non-weight-bearing MRI. Failure to detect meniscal extrusion has hampered development of effective therapies for osteoarthritis (OA) prevention. Weight-Bearing CT (WBCT) has been found to be more sensitive and accurate for other knee OA features, and more accurate assessment of meniscal damage could potentially improve prediction of worsening joint structure and pain.

OBJECTIVE: To assess the rate of detection and severity of meniscal extrusions on visualized on WBCT vs. on MRI in older adults with or at elevated risk for knee OA.

METHODS: Ancillary to MOST, a longitudinal study of knee OA in older Americans, fixed-flexion knee images were acquired using a prototype WBCT scanner. A 3D dataset with an isotropic resolution of 0.37mm was reconstructed from cone beam projections. MRI was acquired using a 1.5T peripheral scanner with participants seated and the knee semi-flexed. Radiologists, blinded to patient identifiers, scored meniscal extrusion severity on each modality (0/1/2/3). Kellgren-Lawrence (KL) grade of knee OA was collected as part of MOST.

RESULTS: Of 864 participants with WBCT imaging of the knees, 284 had MRI read for meniscal extrusion. WBCT detected extrusion not detected on MRI in 27.1% of medial and 8.5% of lateral menisci and higher grades of extrusion for 30.6% of medial and 8.8% of lateral menisci (full results in Tables 1 & 2). Knees with greater medial and lateral extrusions visualized on WBCT were predominantly those with early OA (KL<2 for 80.5% and 64% respectively).

Table 1. Medial Meniscal Extrusion Scores on WBCT and MRI

Extrusion Score on WBCT	Extrusion Score on MRI				
	0	1	2	3	Row Totals
Unreadable	26	8	2	0	36
0	120	16	1	0	137
1	72	17	1	0	90
2	7	8	4	2	21
Column Totals	225	49	8	2	284

Table 2. Lateral Meniscal Extrusion Scores on WBCT and MRI

Extrusion Score on WBCT	Extrusion Score on MRI		
	0	1	Row Totals
Unreadable	1	0	1
0	257	1	258
1	24	0	24
2	0	1	1
Column Totals	282	2	284

CONCLUSION: WBCT detects meniscal extrusions not detected on standard MRI. Detection of this risk factor for OA progression people with early disease supports a need to assess longitudinal associations between meniscal extrusion detected on WBCT and worsening of pain and joint structure.

SPONSOR: NIH-NIAMS R01 AR071648, NIH-NIA U01 AG018832, U01 AG19069
DICLOSURE STATEMENT: AG and FWR are shareholders in BICL, LLC. AG is Consultant to Pfizer, TissueGene, MerckSerono, Regeneron, AstraZeneca and Novartis. FWR is Consultant to Calibr.
CORRESPONDENCE ADDRESS: nsegal@kumc.edu

KNEE OSTEOARTHRITIS PREDICTION CHALLENGE (KNOAP2020): OULU MIPT TEAM SUBMISSIONS

^*Panfilov E., ^*Bayramoglu N., ^*Nguyen H.H., **Nieminen M.T., **Saarakkala S., ***Tiulpin A.

^equal contribution
*University of Oulu, Oulu, Finland
**University of Oulu & Oulu University Hospital, Oulu, Finland
***Aalto University, Espoo, Finland & Ailean Technologies Oy & University of Oulu, Oulu, Finland

INTRODUCTION: Knee OA prediction challenge (KNOAP2020) was organized to objectively compare state-of-the-art methods for prediction of incident symptomatic radiographic knee OA within 78 months (6.5 years). Our team developed a set of predictive models based on the combination of clinical and imaging-derived (both X-ray and MRI) features.

OBJECTIVE: To develop computational models to predict incident symptomatic radiographic knee OA within 78 months (6.5 years) from multi-modal data with good generalization to external sample.

METHODS: Public (npub=30) and private (npr=423) subsets of the PROOF dataset (KNOAP) were provided by the organizers. The public part of the dataset was provided with labels, and the private part was used to make the final submission. For the method development, we used a subset of the OAI dataset with characteristics (age, sex, BMI) matching to the challenge dataset. We used the PA bilateral X-rays (XR), sagittal 3D DESS MRI scans from OAI and sagittal DESS-like scans from KNOAP (MRI), and the clinical features provided by the organizers. Three different models were ensembled using a Naive Bayes classifier (NB). The models were based on: clinical + joint shape and joint space (JS2, Bayramoglu et al., 2020) + MRI-based features (Model A), XR (Model B), clinical + JS2 features (Model C). All the basic models were developed using the same 5-fold cross-validation scheme.
Model A was built using a Gradient Boosting Machine (GBM) and incorporated the considered clinical features (excluding "Varus", k=10), JS2 (k=221), morphological cartilage features (k=86), the same features for the contralateral knee, and the binary indicator of the knee side. Cartilage morphology was quantified by the sub-regional volume and average thickness extracted from the sub-regions defined in Wirth et al., 2008 (Panfilov et al., 2020). GBM hyper-parameters were optimized via a grid search. Model B used only X-ray-based features obtained using a deep convolutional neural network (based on ResNet18 architecture) trained to predict the progression from the PA XR images. Age, sex, BMI, KLG, and JS2 were employed in Model C using a GBM.
Models A, B, and C were trained and evaluated independently. Eventually, the predictions of models A, B, and C were ensembled using a NB with isotonic calibration to form a final submission. The NB parameters were independently optimized on the matched subset of OAI (internal test split) and KNOAP (public split) datasets to produce two ensemble models, corresponding to OuluMIPT-3 and OuluMIPT-5, respectively. Performances of the models were assessed using the area under the ROC curve (ROC AUC) and balanced accuracy (BACC).

RESULTS: The final ensemble method (submission OuluMIPT-3) yielded BACC of 0.698 and ROC AUC of 0.858 on the testing subset of the OAI sample. On the public subset of KNOAP our model obtained a BACC of 0.720, ROC AUC of 0.820. On the preliminary leaderboard, the method was placed 2nd with ROC AUC=0.624(0.546-0.692) and BACC=0.587(0.520-0.648). OuluMIPT-5 submission achieved the 3rd place. Predictions of Models A, B, and C were also submitted independently, but were placed lower on the preliminary leaderboard (9th, 12th, 6th, respectively).

CONCLUSION: Our results suggest that machine learning- and deep learning-based methods may predict combined knee OA progression within 6.5 years from multi-modal data. The ranking of our submissions indicates that ensembling of the diverse models may be beneficial to achieve higher predictive performance. Future work could investigate deeper the ways to improve generalization of the model to new datasets.

SPONSOR: Strategic funding of University of Oulu, Infotech Oulu.
DICLOSURE STATEMENT: A.T. is a co-founder of Ailean Technologies Oy.
ACKNOWLEDGMENT: Claudia Linder for providing BoneFinder software.
CORRESPONDENCE ADDRESS: egor.panfilov@oulu.fi, neslihan.bayramoglu@oulu.fi, huy.nguyen@oulu.fi