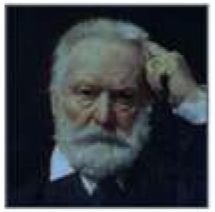
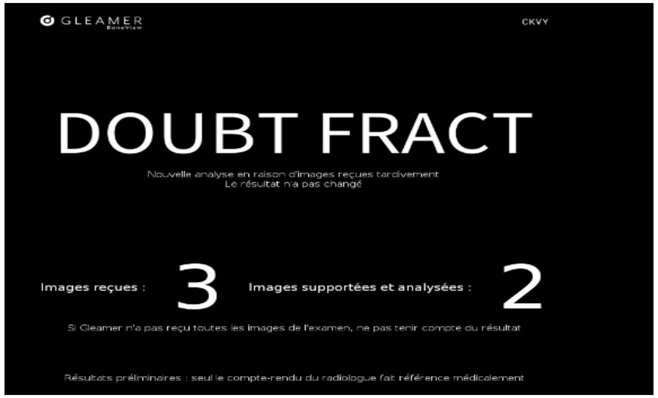
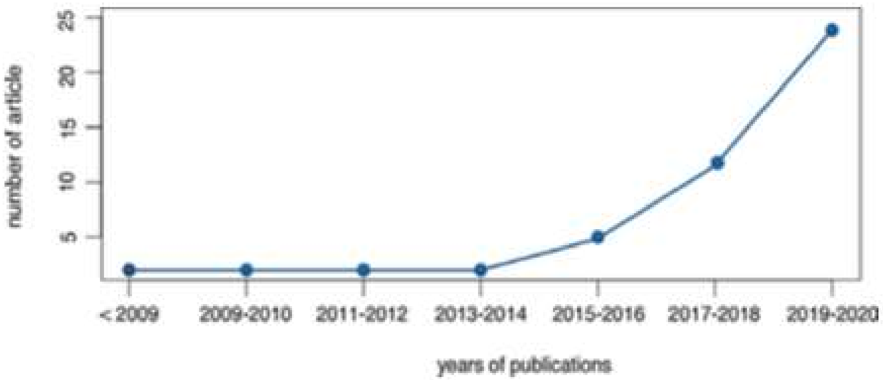
Lecture automatisée d’imagerie en rhumatologie
Rêve ou réalité?



InnoVatIon en ReCherche osTéOaRticulaire de la région HUGO – Réseau VICTOR HUGO





Pr Alain Saraux, Brest

RMD Open
Rheumatic & Musculoskeletal Diseases

ORIGINAL ARTICLE

Multireader assessment as an alternative to reference assessment to improve the detection of radiographic progression in a large longitudinal cohort of rheumatoid arthritis (ESPOIR)

Frederique Gandjbakhch,[1,2] Benjamin Granger,[2,3] Romain Freund,[1,2] Violaine Foltz,[1,2] Sandrine Jousse-Joulin,[4] Valerie Devauchelle,[4] Mona Afshar,[5] Jean David Albert,[6] Florian Bailly,[1,2] Elodie Constant,[7] Lisa Biale,[8] Morgane Milin,[4] Marion Couderc,[9] Delphine Denarie,[7] Anne Fradin,[10] Virginie Martaille,[11] Audrey Pierreisnard,[1,2] Nicolas Poursac,[12] Alain Saraux,[4] Bruno Fautrel[1,2]

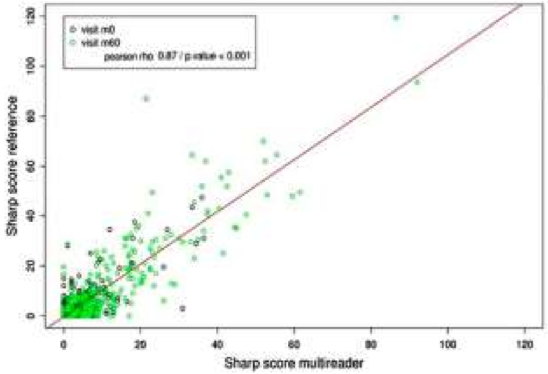


Table 2 Number of patients with structural damage progression (ΔSHS-5 and SRP) in the multireader group and in the reference group

Structural damage progression		Multireader group ΔSHS<5	ΔSHS≥5
Reference group	ΔSHS<5	382	3
	ΔSHS≥5	1	0
		No SRP (ΔSHS<SDC)	**SRP (ΔSHS≥SDC)**
Reference group	No SRP (ΔSHS<SDC)	384	1
	SRP (ΔSHS≥SDC)	1	0

SHS, van der Heijde-modified Sharp score; ΔSHS, SHS change between month 0 (M0) and M60; ΔSHS-5, SHS change between M0 and M60>5; SRP, significant radiographic progression.

- After training, multireader assessment of radiographic structural damage progression is comparable to reference assessment.

Gandjbakhch F, *et al. RMD Open* 2017;**3**:e000343.

Original article

Femoro-tibial knee osteoarthritis: One or two X-rays? Results from a population-based study

Christian-Hubert Roux[a,*], Bernard Mazieres[b], Evelyne Verrouil[b], Anne-Christine Rat[c], Patrice Fardellone[d], Bruno Fautrel[e], Jacques Pouchot[f], Alain Saraux[g], Francis Guillemin[c], Liana Euller-Ziegler[a], Joël Coste[h,1]

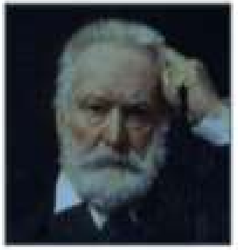
Objective: Our objective was to compare the use of both anteroposterior (AP) extended-knee X-ray and semi-flexed X-ray (current gold standard) versus the use of semi-flexed X-ray alone to detect femoro-tibial osteoarthritis (OA).
Methods: Individuals 40 to 75 years of age with symptomatic hip and/or knee OA (Kellgren/Lawrence [KL] score ≥ 2) were recruited using a multiregional prevalence survey in France. Both AP and schuss X-rays were performed and read; two years later, the same examiner, blinded to the results of the first reading, performed a second reading of the schuss X-ray. We compared the KL stages of each knee and analyzed osteophyte detection and localization, joint space narrowing (JSN), and the relationship to obesity.
Results: The analysis included 350 participants with OA of various stages. Comparing the two readings showed that a higher proportion of patients had KL ≥ 2 when the two X-ray views were combined (right knee: $P < 0.0001$; left knee: $P < 0.001$). There were no differences when using the schuss X-ray alone versus in combination with an AP X-ray in terms of detecting JSN, osteophytes. A comparison of schuss X-ray alone versus AP X-ray alone demonstrated the superiority of the schuss view for evaluating JSN ($P = 0.0001$ and $P = 0.0001$) and no difference in osteophyte detection.
Conclusion: Our study shows that the schuss view alone was sufficient for detecting knee osteophytes and JSN. Using one X-ray rather than two will reduce medical costs and irradiation burden. Using two views seems preferable for epidemiological studies.

Comparison of the schuss view and the anteroposterior (AP) view on conventional X-rays of right and left knees in the 350 osteoarthritis study subjects.

	AP view (KL < 2) *n* (%)	AP view (KL ≥ 2) *n* (%)	Total *n* (%)	*P*-value
Right knee				
Schuss (KL < 2)	245 (70)	22 (6)	267 (76)	
Schuss (KL ≥ 2)	10 (3)	73 (21)	83 (24)	
Total	255 (73)	95 (27)	350 (100)	
Left knee				
Schuss (KL < 2)	262 (75)	24 (7)	286 (82)	
Schuss (KL ≥ 2)	16 (5)	48 (14)	64 (18)	
Total	278 (79)	72 (21)	350 (100)	0.09[a]

KL: Kellgren and Lawrence score.

[a] Comparison schuss versus standard view.

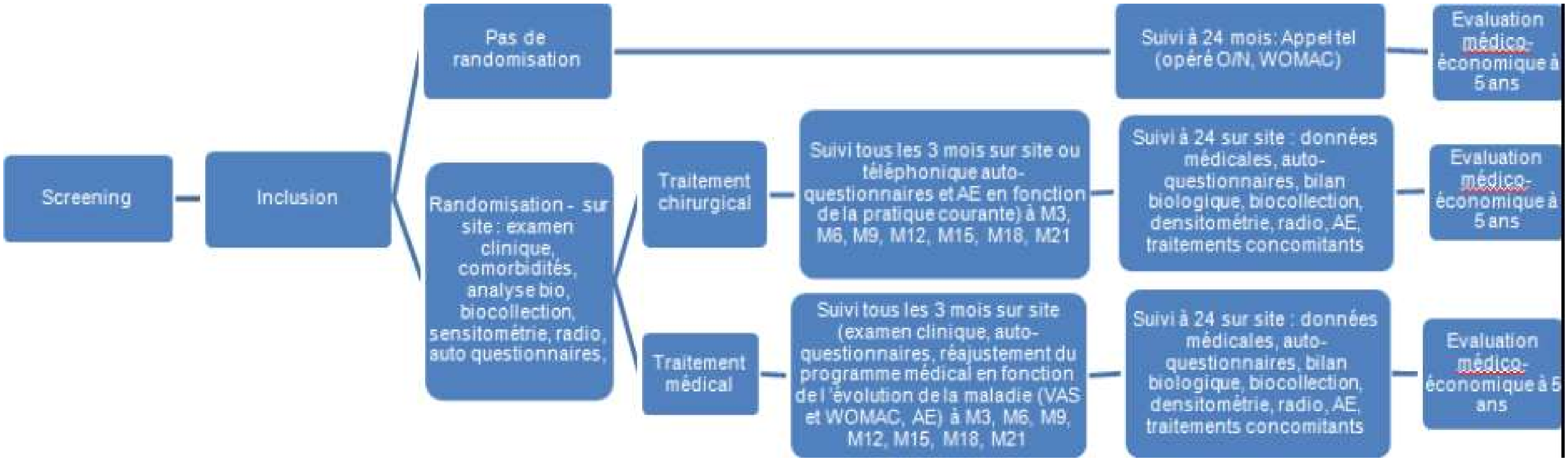


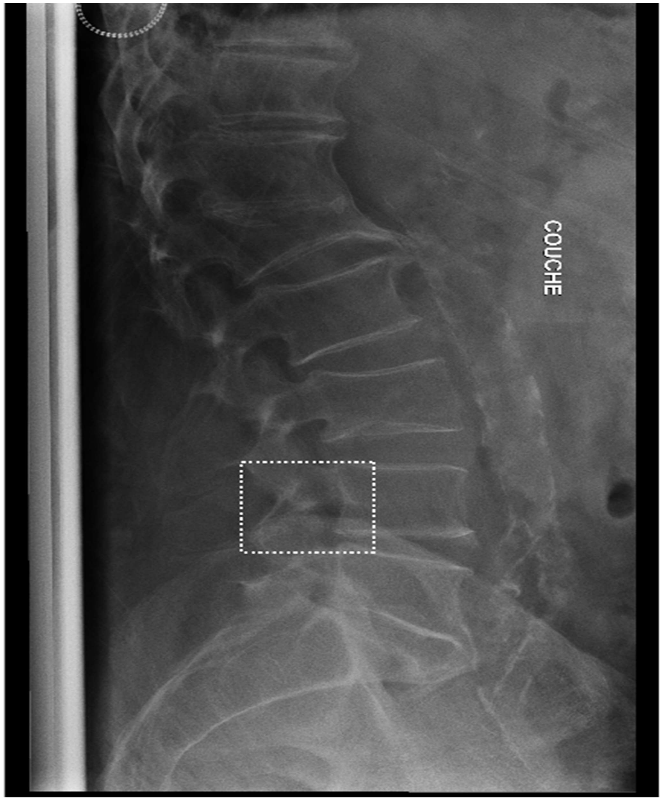
InnoVation en ReCherche osTéOaRticulaire de la région HUGO – Réseau VICTOR HUGO

INCREDIBLE !

Objectif principal

Comparaison du traitement médical multimodal au traitement chirurgical (prothèse) de gonarthroses sévères (radiographique, symptomatique, avec une indication chirurgicale) en termes de douleur (évaluation visuelle analogique de douleur) et de fonction («Western Ontario and McMaster University Osteoarthritis») à 2 ans.



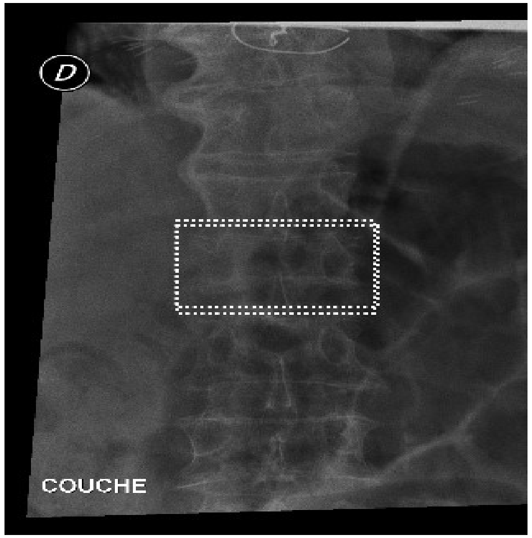


GLEAMER

GLEAMER

NOT AVAILABLE

Partie(s) anatomique(s) "ABDOMEN" non supportée(s)



GLEAMER

GLEAMER

CKVY

DOUBT FRACT

Nouvelle analyse en raison d'images reçues tardivement
Le résultat n'a pas changé

Images reçues : 3 Images supportées et analysées : 2

Si Gleamer n'a pas reçu toutes les images de l'examen, ne pas tenir compte du résultat

Résultats préliminaires : seul le compte-rendu du radiologue fait référence médicalement

Exemple d'IA du quotidien

- ❖ Smartphone (nombreuses fonctions photo, reconnaissance vocale,….)
- ❖ Voiture intelligente (endormissement, aide au parking, navigation…)
- ❖ Media et musique (proposition d'équivalent, reconnaissance…)
- ❖ Jeu vidéo (mouvement selon le jeu….)
- ❖ Annonces et publicités en lignes (cible les personnes)
- ❖ Navigation dans l'espace et en ligne (google map…)
- ❖ Maison connectée (thermostat, vol, chute…)
- ❖ Banque (surveillance, bilan….)
- ❖ Reconnaissance faciale....

A quoi sert l’IA en médecine

- Pour le patient
  - Participation du patient à sa prise en charge (auto-évaluation, PRO, ..)
  - Réponse aux patients par un chatbot (ex: Versus Arthritis, UK Charity, avec IBM Watson)
  - Remplacer la main par la parole ou la parole par la main pour les patients handicapés
  - Aide aux patient handicapés (ouverture de porte, exosquelettes,….)
- Pour le médecin
  - Sécurité par warning (interaction médicamenteuse, détection d’anomalie d’imagerie ou lame anapath)
  - Proposition d’interprétation d’images (Stanford University’s CheXNet au moins égal au radiologue pour le diagnostic de pneumonie, fond d’oeil, mélanome…)
  - Reconstruction d’images
  - Diagnostic et prédiction en médecine à partir de nombreuses données hétérogènes/Médecine de précision
  - Formation et revue de littérature
  - Précision de l’acte (Robot)
  - Choix de l’outil en chirurgie ou NRI
- Pour les établissements
  - Garder les professionnels et les patients (robot chirurgical)
  - Fouille de données (recherche de candidats à un essai thérapeutique)
  - Prédire les flux aux urgences
  - Traitement de tâches répétitives, à faible valeur ajoutée (codage par exemple)

L’intelligence artificielle en pratique et sa place dans l’imagerie

- Elle peut utiliser
  - Computer vision (CV): Le « convolutional neuronal networks » (CNN, réseau neuronal convolutif) permet d’obtenir des algorithmes de vision spatiale (IRM, anapath….)
  - « Natural language processing » (NLP, traitement automatique du langage naturel): Le « recurrent neural networks » (réseau de neurones récurrents) permet d’obtenir des algorithmes à partir de texte (données de santé numérique)
  - « Reinforcement learning » (RL, apprentissage par renforcement): Il apprend à partir des erreurs à faire des corrections simples (par exemple un geste avec un robot)

Lecture automatisée d’images (fond d’œil, lésion cutanée, …)

- **L'intelligence artificielle peut ainsi être utilisée pour des tâches d'assistance technique.**
  - particulièrement efficace pour faire de l'analyse d'image
  - via l'utilisation de réseaux de neurones à convolutions (CNN)
  - utilisée dans le cadre de l'imagerie par rayons X, ultrasons, IRM, anatomopathologie pour identifier des structures particulières au sein d'images
  - voire pour fournir directement une aide au diagnostic (pneumonies, atteinte rétinienne des diabétiques, suspicion de mélanome…).

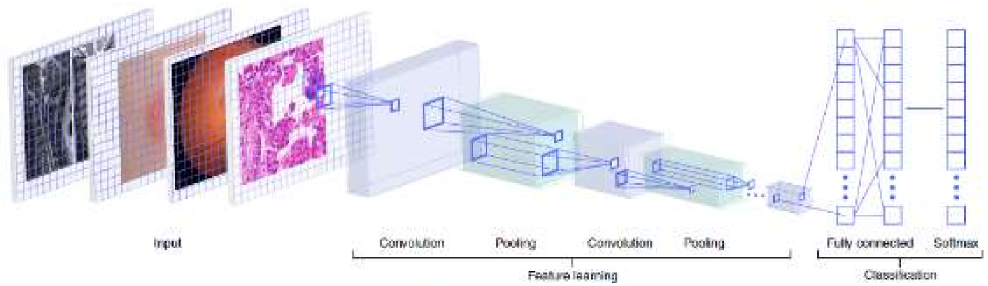


Fig. 2 | Medical imaging. CNNs can be trained on a variety of medical imagery, including radiology, pathology, dermatology, and ophthalmology. Information flows left to right. CNNs take input images and sequentially transform them, using simple operations such as convolutional, pooling, and fully connected layers, into flattened vectors. The elements of the output vector (softmax layer) represent the probabilities of the presence of disease. During the training process, the internal parameters of the network layers are iteratively adjusted to improve accuracy. Typically, lower layers (left) learn simple image features—edges and basic shapes—which influence the high-level representations (right). Prediction tasks include both classification of the images (i.e., cancerous versus benign) as well as localization of medical features such as tumors.

Recherche sur pubmed le 6 octobre 2022

- artificial intelligence radiology bone
  - 1496 résultats
- artificial intelligence radiology joint
  - 681 resultats
- artificial intelligence radiology spine
  - 468 resultats
- Artificial intelligence radiomics
  - 1783

Artificial Intelligence and Deep Learning for Rheumatologists: A Primer and Review of the Literature

Christopher McMaster [1, 2, 3, 4], Alix Bird [5], David Fl Liew [1, 2, 6], Russell R Buchanan [1, 6], Claire E Owen [1, 6], Wendy W Chapman [3], Douglas Ev Pires [3, 4]Arthritis Rheumatol . 2022 Jul 20.

Problem	Data type	Model	Implications
Identifying giant cell arteritis features from temporal artery biopsy reports (7)	Text	Transformer	Accurate auditing of temporal artery biopsy reports can be performed using deep learning, although this performance drops when tested across centers.
Classification of Hep-2 cells based on ANA-IIF patterns (8)	Images	CNN	Automated ANA classification based on Hep-2 cells is approaching expert human performance.
EULAR-OMERACT Synovitis Scoring (OESS) from synovial ultrasound (9)	Images	CNN	Deep learning can identify synovitis on ultrasound with a high degree of accuracy.
Sharp/van der Heijde scoring using hand & foot radiographs (10)	Images	CNN	Radiographic scoring for rheumatoid arthritis is improving, but still requires work for clinical implementation.
Predicting progression (any increase in KL score) of knee OA based on baseline knee radiograph plus other clinical features (11)	Images	CNN	Radiographic progression in knee OA can be predicted with a combination of clinical features and baseline x-ray using deep learning, however there are definitely unmeasured factors we are missing in these models.
Identifying 'halo sign' on temporal artery ultrasound images (12)	Images	CNN	Deep learning has significant potential for automated identification of the halo sign, however ensuring standardized image acquisition is a major barrier to implementation.
Predicting future rheumatoid arthritis disease activity (controlled vs. uncontrolled) using clinical data from previous encounters (13)	Electronic Health Records	RNN	Deep learning can predict future disease activity from past disease activity and baseline factors, however performance drops significantly when this model is tested at a second center, suggesting that there is substantial heterogeneity between centers that must be accounted for in future models.

Table 1: Current applications of deep learning in rheumatology.

Artificial Intelligence in Fracture Detection:

A Systematic Review and Meta-Analysis

*Rachel Y. L. Kuo, MB BChir, MA, MRCS • Conrad Harrison, BSc, MBBS, MRCS •
Terry-Ann Curran, MB BCh BAO, MD • Benjamin Jones, BMBCh, BA •
Alexander Freethy, BSc, MBBS, MSc, MRCS • David Cussons, BSc, MBBS • Max Stewart, MB BChir, BA •
Gary S. Collins, BSc, PhD • Dominic Furniss, DM, MA, MBBCh, FRCS (Plast)*

Table 4: Pooled Sensitivities, Specificities, and Areas Under the Curve for Artificial Intelligence Algorithms and Clinicians

Parameter	Sensitivity (%)	Specificity (%)	AUC	No. of Contingency Tables
Algorithms, internal validation, all studies	92 (88, 94)	91 (88, 93)	0.97 (0.95, 0.98)	37
Studies with low bias	90 (86, 93)	89 (85, 92)	0.95 (0.93, 0.97)	21
Clinicians, internal validation, all studies	91 (85, 95)	92 (89, 95)	0.97 (0.95, 0.98)	36
Studies with low bias	89 (76, 95)	86 (80, 90)	0.93 (0.90, 0.95)	13
Algorithms, external validation, all studies	91 (84, 85)	91 (81, 95)	0.96 (0.94, 0.98)	15
Studies with low bias	89 (76, 95)	80 (74, 85)	0.87 (0.84, 0.90)	10
Clinicians, external validation, all studies	94 (90, 96)	94 (91, 95)	0.98 (0.96, 0.99)	23
Studies with low bias	93 (87, 96)	93 (89, 95)	0.97 (0.95, 0.98)	16
Clinicians with AI assistance, all studies	97 (83, 99)	92 (88, 95)	0.95 (0.92, 0.96)	4
Studies with low bias	97 (83, 99)	92 (88, 95)	0.95 (0.92, 0.96)	4

Note.—Data in parentheses are 95% CIs. Results of all studies and studies with low bias are compared. AI = artificial intelligence, AUC = area under the receiver operating characteristic curve.

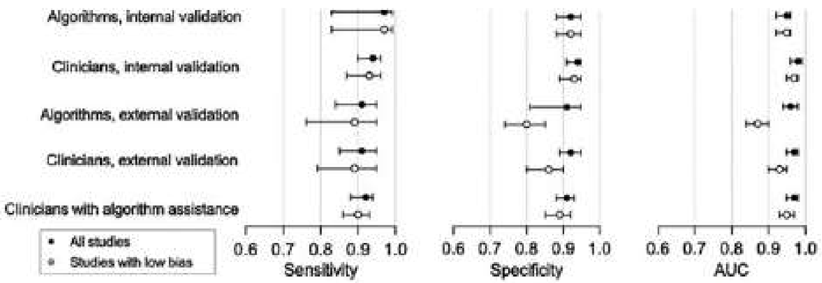


Figure 6: Summary of pooled sensitivity, specificity, and area under the curve (AUC) of algorithms and clinicians comparing all studies versus low-bias studies with 95% CIs.

Lecture radiographique de fractures

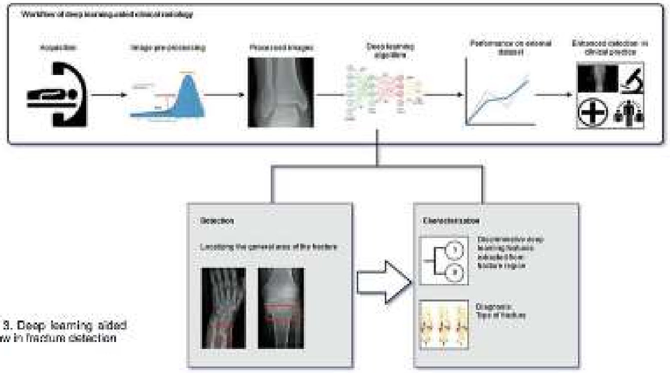


Figure 3. Deep learning aided workflow in fracture detection

Acta Orthopaedica 2020; 91 (2): 215–220

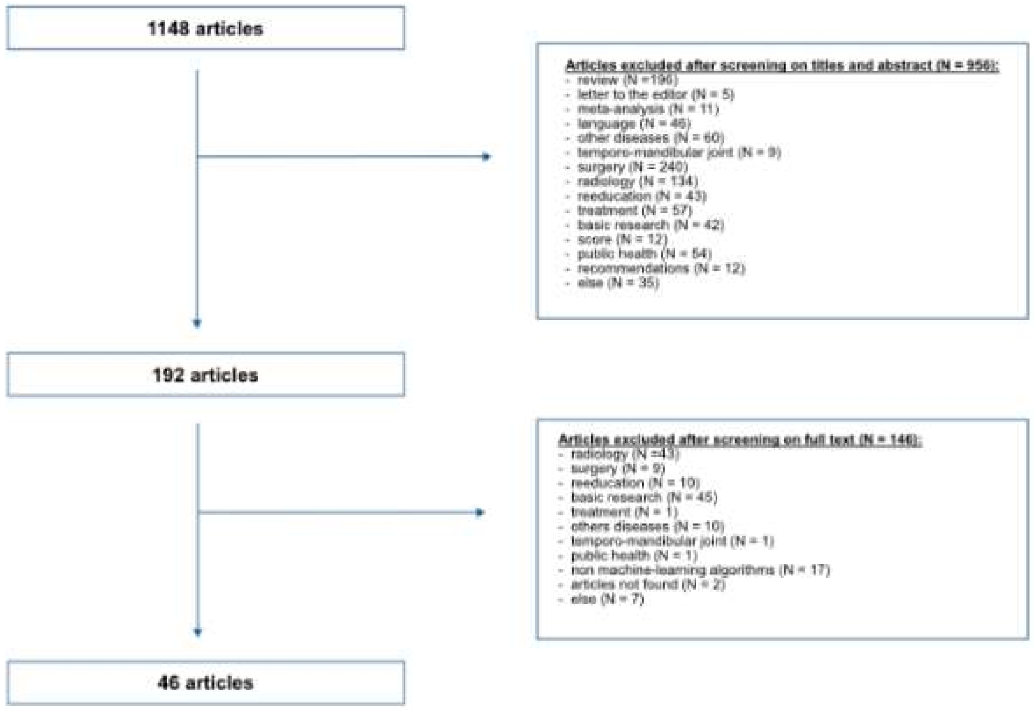
Deep learning in fracture detection: a narrative review

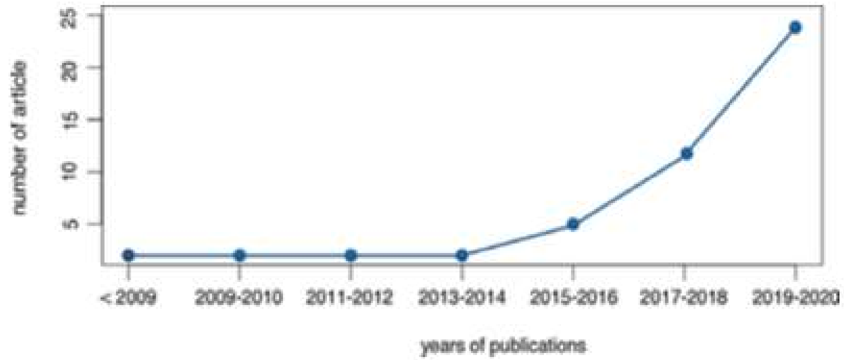
Pishtiwan H S KALMET [1]*, Sebastian SANDULEANU [2]*, Sergey PRIMAKOV [2], Guangyao WU [2], Arthur JOCHEMS [2], Turkey REFAEE [2], Abdalla IBRAHIM [2], Luca v. HULST [1], Philippe LAMBIN [2], and Martijn POEZE [1,3]

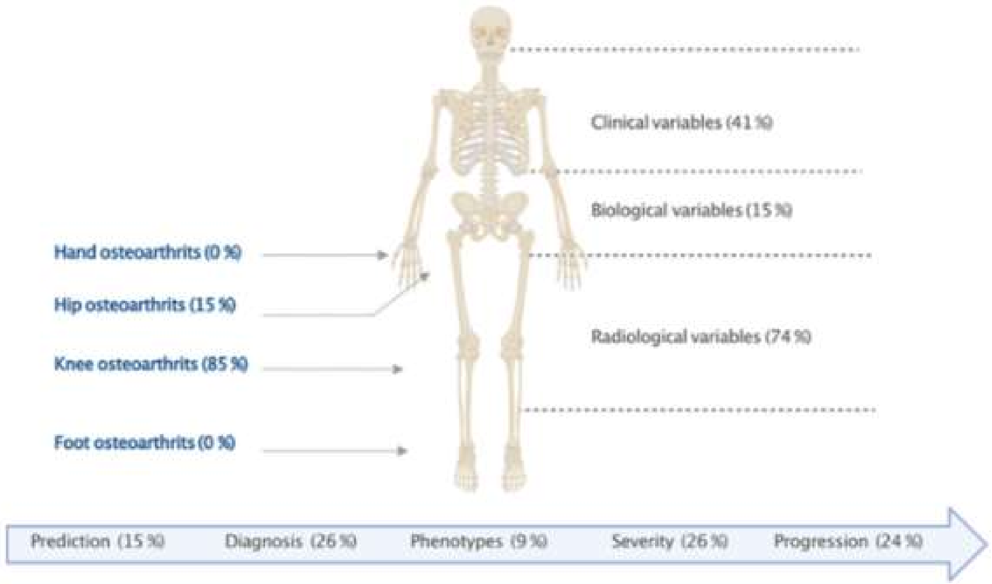
Reference	Region of interest	Modality	Conclusion	Performance (metric)
Olczak et al. 2017	Wrist/Hand/Ankle	Radiographs	This study supports the use of orthopaedic radiographs of artificial intelligence, which can perform at a human level	0.83 (accuracy)
Kim et al. 2018	Wrist	Radiographs	The AUC scores for this test were comparable tostate-of-the-art providing proof of concept for transfer learning from CNNs in fracture detection on plain radiographs	0.95 (AUC) 0.90 (sensitivity) 0.88 (specificity)
Chung et al. 2018	Proximal humerus	Radiographs	The use of artificial intelligence can accurately detect and classify proximal humerus fractures on plain shoulder AP radiographs	Detection: 0.96 (accuracy) 1 (AUC) 0.99 (sensitivity) 0.97 (specificity) Classification: 0.65–0.86 (accuracy) 0.90–0.98 (AUC) 0.88–0.97 (sensitivity) 0.83–0.94 (specificity)
Heimer et al. 2018	Skull	CT	Classification based on the existence of skull fractures on CMIPs with deep learning is feasible	0.97 (AUC) 0.91 (sensitivity) 0.88 (specificity)
Lindsey et al. 2018	Wrist	Radiographs	Deep learning methods are a mechanism by which senior medical specialists can deliver their expertise to generalists on the front lines of medicine, thereby providing substantial improvements to patient care	0.97 (AUC) on Test set1 0.98 (AUC) on Test set2
Tomita et al. 2018	Pelvis	CT	The proposed system will assist and improve OVF diagnosis in clinical settings by pre-screening routine CT examinations and flagging suspicious cases prior to review by radiologists	0.89 (accuracy) 0.91 (F1 score)
Pranata et al. 2019	Calcaneus	CT	The feasibility using deep CNN and SURF for computer-aided classification and detection of the location of calcaneus fractures in CT images	0.98 (accuracy)
Adams et al. 2019	Pelvis	Radiographs	As impressive as recognising fractures is for a DCNN, similar learning can be achieved by top-performing medically naïve humans with less than 1 hour of perceptual training	0.91 (accuracy) 0.98 (AUC)

Use of machine learning in osteoarthritis research: a systematic literature review

Marie Binvignat [1,2,3] Valentina Pedoia,[4] Atul J Butte,[2] Karine Louati,[1] David Klatzmann,[3,5] Francis Berenbaum [1] Encarnita Mariotti-Ferrandiz,[3] Jérémie Sellam[1]







Binvignat M, *et al. RMD Open* 2022;**8**:e001998.

Artificial intelligence in diagnosis of knee osteoarthritis and prediction of arthroplasty outcomes: a review

Lee *et al. Arthroplasty* (2022) 4:16
https://doi.org/10.1186/s42836-022-00118-7

Lok Sze Lee[1], Ping Keung Chan[1*], Chunyi Wen[2], Wing Chiu Fung[1], Amy Cheung[3], Vincent Wai Kwan Chan[3], Man Hong Cheung[1], Henry Fu[1], Chun Hoi Yan[4] and Kwong Yuen Chiu[1]

OA diagnosis and TKA need	El-Galaly, A.	2020	Clinical Orthopaedics and Related Research
	Heisinger, S.	2020	Journal of Clinical Medicine
	Jafarzadeh, S.	2020	Osteoarthritis Cartilage
	Leung, K.	2020	Radiology
	Tolpadi, A.A.	2020	Scientific Reports
	Yi, P. H.	2020	Knee
	Norman, B.	2019	Journal of Digital Imaging
	Tiulpin, A.	2018	Scientific reports

Methods: PubMed and EMBASE databases were searched for articles published in peer-reviewed journals between January 1, 2010 and May 31, 2021. The terms included: 'artificial intelligence', 'machine learning', 'knee', 'osteoarthritis', and 'arthroplasty'. We selected studies focusing on the use of AI in diagnosis of knee osteoarthritis, prediction of the need for total knee arthroplasty, and prediction of outcomes of total knee arthroplasty. Non-English language articles and articles with no English translation were excluded. A reviewer screened the articles for the relevance to the research questions and strength of evidence.

Author (Year)	Journal	Prediction outcome	AI/ML algorithm(s)	Statistical performance	Strengths	Weaknesses	Clinical significance of study
Norman (2019) [18]	Journal of Digital Imaging	OA severity (KL grade)	DenseNet neural network architectures	Sensitivity & specificity: 84% & 86% (KL grades 0–1), 70% & 84% (KL grade 2), 69% & 97% (KL grade 3), 86% & 99% (KL grade 4).	Comparable sensitivity and specificity to manual KL grading and previous automatic systems employing different AI/ML algorithms	Training, validation and testing sets were selected from the same dataset. Misclassifications of KL grading typically occurred when there was hardware in the knee.	Provides additional data supporting the potential of AI in automatic assessment of OA radiological severity.
Tiulpin (2018) [19]	Scientific reports	OA severity (KL grade)	Deep Siamese CNN architecture	Average multi-class accuracy: 66.71%. AUC: 0.93. Kappa coefficient (agreement with expert annotations on test dataset): 0.83 (excellent). MSE value: 0.48.	Different datasets used for initial training and testing	Validation and testing sets were selected from the same dataset.	The provision of probability distributions for each KL grade prediction may assist clinicians in choosing KL grade in ambiguous cases.
Heisinger (2020) [13]	Journal of Clinical Medicine	Need for TKA	Artificial neural networks (ANNs) with linear, radial basis function and three-layer perceptron neural networks architectures	Total percentage of correctly predicted knees: 80%. Positive predictive value: 84%. Negative predictive value: 73%. Sensitivity: 41%. Specificity 30%.	First study to consider longitudinal change in symptomology (pain, function, quality of life) and radiographic structural change in a 4-year period prior to TKA	Training and testing sets were selected from the same dataset.	Future externally validated algorithms that can predict TKA need in advance using routinely available patient data could be highly useful for decisions for referral and triage in a primary care setting.
Leung (2020) [15]	Radiology	Need for TKA	Multitask deep learning model (ResNet34) trained with transfer learning	AUC: 0.87. Sensitivity: 83%. Specificity: 77%.	First study to directly predict TKA from knee radiographs using deep learning model	Limited data size (radiographs from 728 individuals in total) / Training and testing sets were selected from the same dataset.	TKA prediction models solely based on radiological data have limited clinical utility, although they may serve as a reference for future ML studies.
El-Galaly (2020) [12]	Clinical Orthopaedics and Related Research	Need for early revision TKA	LASSO regression, random forest classifier, gradient boosting model, neural network	AUCs: 0.57–0.60.	First study to predict early revision TKA ($\leq$ 2 years of primary TKA) using preoperative patient data from arthroplasty registries / Temporal external validation was conducted (testing set selected from a separate hold-out year not included in training set).	Training and testing sets were selected from the same dataset.	Results from this study suggest that future models predicting early revision TKA may benefit from including more pre-operative information or predicting revision over a longer follow-up duration.

Lecture radiographique d’arthrose

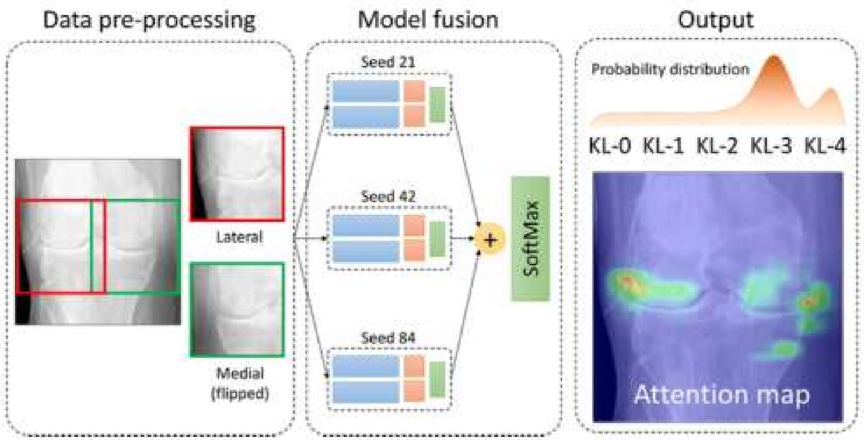
SCIENTIFIC REPORTS

OPEN

Automatic Knee Osteoarthritis Diagnosis from Plain Radiographs: A Deep Learning-Based Approach

Received: 21 July 2017

Aleksei Tiulpin[1], Jérôme Thevenot[1], Esa Rahtu[4], Petri Lehenkari[2] & Simo Saarakkala[1,3]



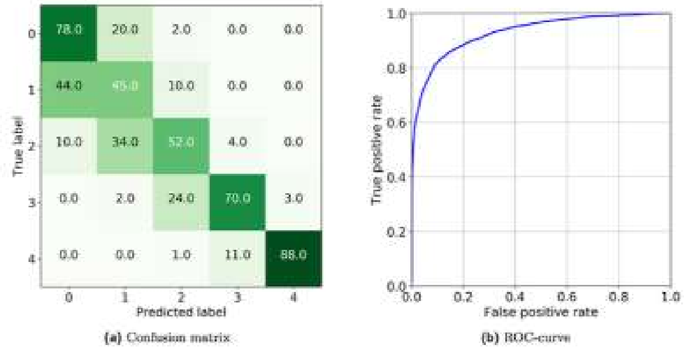


Figure 4. (a) Confusion matrix of KL grading and (b) ROC curve for radiographic OA diagnosis KL ≥ 2 produced using our method. Average multi-class accuracy is 66.71%, and AUC value is 0.93. Corresponding Kappa coefficient and MSE value are 0.83 and 0.48, respectively.

SCIENTIFIC REPORTS | (2018) 8:1727 | DOI:10.1038/s41598-018-20132-7

RESEARCH ARTICLE Open Access

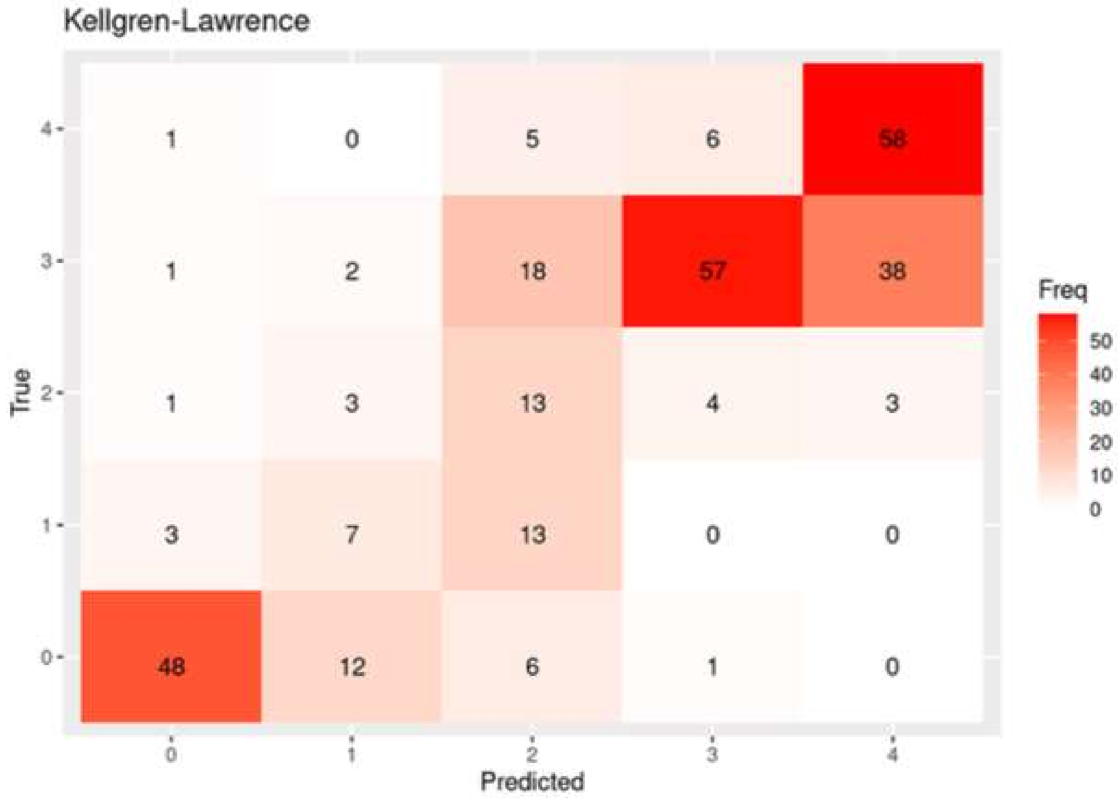
Automating classification of osteoarthritis according to Kellgren-Lawrence in the knee using deep learning in an unfiltered adult population

Simon Olsson, Ehsan Akbarian, Anna Lind, Ali Sharif Razavian and Max Gordon*

Table 3 Network performance on outcome measures. Kellgren & Lawrence grades are displayed separately and merged

	Cases ($n = 300$)	Sensitivity (%)	Specificity (%)	Youden's J	AUC (95% CI)
Kellgren-Lawrence					
0	67	97	88	0.85	0.97 (0.93 to 0.99)
1	23	96	75	0.70	0.88 (0.83 to 0.92)
2	24	92	61	0.53	0.80 (0.73 to 0.86)
3	116	92	71	0.63	0.87 (0.83 to 0.90)
4	70	84	78	0.63	0.87 (0.83 to 0.91)
Grouped Kellgren-Lawrence					
0 to 1	90	88	95	0.83	0.96 (0.94 to 0.98)
0 to 2	114	83	97	0.81	0.97 (0.95 to 0.98)
1 to 3	163	80	74	0.53	0.82 (0.77 to 0.87)
2 to 4	210	97	83	0.80	0.96 (0.94 to 0.98)
3 to 4	186	96	88	0.84	0.97 (0.95 to 0.99)

Methods: We selected 6103 radiographic exams of the knee taken at Danderyd University Hospital between the years 2002-2016 and manually categorized them according to the Kellgren & Lawrence grading scale (KL). We then trained a convolutional neural network (CNN) of ResNet architecture using PyTorch. We evaluated the results against a test set of 300 exams that had been reviewed independently by two senior orthopedic surgeons who settled eventual interobserver disagreements through consensus sessions.



Evaluation of artificial intelligence models for osteoarthritis of the knee using deep learning algorithms for orthopedic radiographs

Anjali Tiwari, Murali Poduval, Vaibhav Bagaria

Table 1 Data splits in training, testing and validation subsets according to Kellgren-Lawrence grades

Osteoarthritis Kellgren-Lawrence grade	Training		Testing		Validation	
	Samples, *n*	Proportion, %	Samples, *n*	Proportion, %	Samples, *n*	Proportion, %
0	255	17.6	37	17.6	73	17.7
1	213	14.7	31	14.7	61	14.8
2	164	11.4	24	11.4	47	11.4
3	237	16.4	35	16.7	68	16.4
4	576	39.9	83	39.5	164	39.7
Total	1445	100	210	100	413	100

World J Orthop 2022 June 18; 13(6): 603-614

Parameter	
Accuracy	Determines the accuracy of the standalone model inaccuracy to detect the presence of KOA and its classification in the input image
Precision	True positive/true positive + false positive
Recall	True positive/true positive + false negative
Loss	Determines the loss of the model

Table 3 Performance comparison of various transfer learning convolutional neural network models and eight expert human interpretations used for the development of deep learning algorithm for orthopedic radiographs

Model name	Accuracy	Precision	Recall	Loss	Outcome
ResNet50	54.29%	61.03%	39.52%	1.06	Average
VGG-16	56.68%	67.56%	35.02%	1.10	Average
InceptionV3	87.34%	89.19%	85.67%	0.35	Good
MobilnetV2	82.15%	84.66%	80.21%	0.46	Average
EfficientnetB7	56.61%	70.09%	38.27%	0.98	Average
DenseNet201	92.87%	93.69%	92.53%	0.20	Best
Xception	82.81%	85.03%	77.05%	0.50	Average
NasNetMobile	80.90%	83.98%	77.30%	0.50	Average
Surgeon	74.22%	79.50%	50.00%	0.25	Good

Conclusion: Cycle du "hype"