Published in final edited form as:

Expert Rev Precis Med Drug Dev. 2021 ; 6(2): 83–85. doi:10.1080/23808993.2021.1848424.

How feasible is the stratification of osteoarthritis phenotypes by means of artificial intelligence?

Amanda E. Nelson[*]

Department of Medicine, Division of Rheumatology, Allergy, and Immunology, Director, Phenotyping and Precision Medicine Resource Core of the UNC Core Center for Clinical Research, University of North Carolina at Chapel Hill School of Medicine, 3300 Doc J. Thurston Building, Campus Box #7280, Chapel Hill, NC, USA, 27599-7280

Keywords

machine learning; phenotype; precision medicine; osteoarthritis

Introduction

Osteoarthritis (OA) is a common and serious disease that involves all of the tissues of an affected joint (e.g., cartilage, bone, meniscus, tendon/ligament, synovium) and can affect one or multiple joints in an individual person, most often the finger joints, knees, hips, and spine [1]. OA is a major and growing contributor to disability worldwide, and is associated with increased comorbidity and excess mortality [1]. Management of OA is focused on modestly effective lifestyle/behavioral interventions such as increased physical activity and weight loss, with pharmacologic therapies directed toward temporary symptomatic relief [2]. Although many clinical trials have been conducted, there are still no effective disease modifying therapies, no proven way to prevent progression, and no cure. This is at least in part due to the lack of appreciation of, and accounting for, the heterogeneity of this complex disease in trials to date [3]. In general, most trials have enrolled all individuals with knee OA defined as the presence of symptoms (e.g., pain, aching, and stiffness) and moderate to severe radiographic change (e.g., osteophytes or joint space narrowing) in at least one knee. This does not account for the diverse mechanisms of disease development, which can be due to mechanical dysfunction, prior injury, metabolic factors, inflammation, or combinations of these. Nor does it address the diversity of presentations, burden of disease (i.e., number/severity of involved joints), chronicity, or numerous other aspects of the disease process in a given individual that may subsequently affect their response to the proposed therapy. This

[*]**Corresponding author:** Amanda E. Nelson, Department of Medicine, Division of Rheumatology, Allergy, and Immunology, Director, Phenotyping and Precision Medicine Resource Core of the UNC Core Center for Clinical Research, University of North Carolina at Chapel Hill School of Medicine, 3300 Doc J. Thurston Building, Campus Box #7280, Chapel Hill, NC, USA, 27599-7280, (919) 966-0553, aenelson@med.unc.edu.

Declaration of interest

The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

Reviewers Disclosure

Peer reviewers on this manuscript have no relevant financial relationships or otherwise to disclose.

brief editorial review seeks to summarize recent work in the area of machine learning and osteoarthritis phenotyping.

Reviews of OA phenotypes and endotypes

OA has been recognized clinically to be heterogeneous for many years, but focused research into potential subgroups has been a more recent development. In a review of this literature in 2016, there was evidence for six potential clinical phenotypes: chronic pain with central sensitization; inflammatory; metabolic syndrome; bone and cartilage metabolism; mechanical overload; and minimal joint disease [4]. These investigators subsequently explored the prevalence of these phenotypes in the Foundation for the National Institutes of Health OA Biomarkers Consortium (OAI/FNIH) dataset, finding representation of all six, as well as a subgroup with substantial overlap and another that was not classifiable as any of these phenotypes [5]. Another review focused on molecular/mechanistic endotypes, particularly those related to inflammation (characterized by, for example, c-reactive protein and interleukin-6), bone (e.g., c-terminal telopeptide of collagen I [CTX-I]), metabolic syndrome (e.g., adipokines, advanced glycation end products), and aging (markers of senescence) [6]. There has been a recent effort to standardize the conduct and reporting of phenotype studies in OA [7], in which phenotypes were noted to be “subtypes of OA that share distinct underlying pathobiological and pain mechanisms and their structural and functional consequences,” that should differ from others in disease-driving factors and/or outcomes.

Artificial intelligence for imaging in OA

To date, the uptake of artificial intelligence (AI, or the use of computer-based intelligence) methodologies in OA research has been predominantly in the area of image analysis. These studies have utilized various AI methods for analyzing radiographic or magnetic resonance (MR) images, in some cases to increase efficiency (as reading these images is currently quite time-consuming and reader dependent) and in others to identify novel features that may define phenotypes.

Conventional radiography

Deep learning (DL) via convolutional neural networks (CNN) has been applied to automatically score severity of knee OA according to the Kellgren-Lawrence grade (KLG), using data from the Multicenter Osteoarthritis (MOST) study (training set) and the Osteoarthritis Initiative (OAI, independent test set). The area under the receiver-operating characteristic curve (AUC) for this method to diagnose radiographic OA was 0.93; the output also included probability distributions for the KLG, framing the level of variation or uncertainty [8]. DL models can also improve diagnostic performance for predicting medial joint space loss from baseline radiographs [9]. A combined joint training model, incorporating demographic and radiographic risk factors along with DL analysis of baseline knee x-rays had an AUC of 0.86, was an improvement over either model alone (AUC=0.66 for traditional; AUC=0.80 for DL) [9]. Both studies also utilized maps (intention or saliency maps) to show the areas where the AI model is focusing to make determinations regarding class membership, improving interpretability [8,9]. Another group utilized machine learning

(ML) with raw radiographic image data (via CNN) as well as medical, examination, and anthropometric data from the OAI (training) and MOST (testing) studies, to predict structural OA progression (defined as increased KLG or total knee replacement [TKR]) [10]. Compared with regression methods (AUC=0.75), CNN of radiographic data alone had better prediction (AUC=0.79), but incorporating the CNN data with clinical and KLG data using gradient boosting machine methodology produced the highest AUC (0.81, 95% confidence interval [CI]: 0.79–0.82) and the greatest precision (0.70, 0.68–0.72) [10].

Magnetic resonance imaging

One small study (n=175) assessed the feasibility of a DL/CNN approach to detect MR-based cartilage lesions, and found sensitivity and specificity comparable to radiologists at various levels of training (around 85%), as well as high diagnostic accuracy (AUC>0.9) compared with an expert musculoskeletal radiologist as the gold standard [11]. Another study utilized baseline T2 relaxation time maps for OAI participants, quantifying differences between subjects with and without radiographic OA using voxel-based relaxometry, followed by diagnosis of OA based on T2 data from a densely connected CNN. This approach was compared with traditional feature extraction and random forest modeling, finding improved sensitivity and specificity for the CNN-based method (77% and 78%, respectively; AUC=0.82) compared with random forest plus principal components from the T2 maps (67% and 72%, respectively; AUC=0.78) [12].

A DL-based pipeline was developed to incorporate imaging (both radiography and MR), clinical, and demographic information to predict eventual TKR in the OAI. This study compared results for the imaging CNN alone, the output of a single regression model using only the non-imaging variables, and an integrated model incorporating all of these data. For knees at the extremes, either with no OA or severe OA at baseline, models using MR data had improved sensitivity compared with those using radiographic data, although models using radiography had superior accuracy. These methods additionally identified potentially novel tissues (i.e., the medial patellar retinaculum, gastrocnemius tendon, and plantaris muscle) as being key for TKR prediction; however, commonly used imaging biomarkers such as cartilage, bone, menisci, and ACL, were not. The final accuracy of the MR-based model was 79%, with sensitivity of 82% and specificity of 78% [13].

Artificial intelligence for clinical phenotyping in OA

We utilized ML of clinical and imaging features to identify groups that did and did not progress by pain and radiographic measures in the OAI/FNIH data set [14]. We employed novel methods specifically for high dimension low sample size settings to contrast these groups, including Distance Weighted Discrimination (a linear discriminant analysis method) and Direction-Projection-Permutation hypothesis testing. Both quantitative and semi-quantitative MR variables were important features to differentiate progressors from non-progressors (z-scores 10.3–11.6), while demographic/clinical and biochemical biomarkers were not as useful (z-scores 1.5–2.4). We were also able to identify the features with the greatest contribution to the identified differences [14]. Additionally, a data-driven

visualization method (t-distributed stochastic neighbor embedding [t-SNE]) provided new insights into the magnitude of sex differences in OA [14].

The experience of pain in OA is itself quite heterogeneous, including aspects of local and central sensitization, making clear the need for phenotypes of pain in OA [15]. In addition, the transitions among pain states are variable and have been described as phenotypes of pain susceptibility. To address this, investigators using data from the MOST study incorporated a variety of self-reported measures of symptoms, psychosocial function, and quantitative sensory testing (QST) results based on an outcome of persistent pain 2 years later. They found four latent classes based on the QST measures (pressure pain sensitivity and temporal summation), with none of the other variables informing the phenotypes. The subgroup with the highest proportion of pressure pain sensitivity and facilitated temporal summation had twice the odds of developing persistent knee pain compared to those with neither; importantly there was no difference between the groups based on radiographic OA [15]. The ability to identify those at risk for persistent or chronic pain could open new opportunities to prevent this morbid transition.

Summary and future directions

OA is a prevalent and heterogeneous disease that is a prime candidate for phenotyping; AI methodologies are just beginning to be employed in this area. There are challenges remaining, including that of missing data (observations and features with missing data are often deleted for these analyses) and external validation (difficult due to a lack of multiple independent cohorts with comparable data). Although the results to date are relatively predictable (in that use of more complex data objects provides improved accuracy in prediction models) and are of modest magnitude (e.g., from an AUC of 0.75 to 0.81), the potential of these approaches to define higher risk groups and direct potential targeted therapies in future clinical trials is apparent. In light of the lack of effective therapies for OA, which can at least in part be ascribed to a lack of targeted intervention trials, phenotyping approaches are crucial for the success of future studies. One early example of the potential of relatively simple subgrouping in OA is a currently active clinical trial of a pharmacologic agent for OA (NCT03928184; clinicaltrials.gov). The eligibility requirements for the trial include not only the usual KLG 2–3, but also places a minimum and maximum on the range for radiographic joint space width as well as self-reported pain and function. It is likely that further progress in the areas of OA phenotyping, endotyping and precision medicine [16], utilizing various aspects of AI, will more dramatically impact the design and implementation of clinical studies in the near future.

Funding

This work was supported in part by: National Institutes of Health / National Institute of Arthritis and Musculoskeletal and Skin Diseases P30AR072580 and R21AR074685.

References

Papers of special note have been highlighted as:

* of interest

** of considerable interest

1. Osteoarthritis Research Society International. Osteoarthritis: A Serious Disease2016 [cited 1–103 p.]. https://www.oarsi.org/sites/default/files/docs/2016/oarsi_white_paper_oa_serious_disease_121416_1.pdf
2. Kolasinski SL, Neogi T, Hochberg MC, et al. 2019 American College of Rheumatology/Arthritis Foundation Guideline for the Management of Osteoarthritis of the Hand, Hip, and Knee. Arthritis Rheumatol. 2020 2;72(2):220–233. [PubMed: 31908163]
3. Deveza LA, Nelson AE, Loeser RF. Phenotypes of osteoarthritis: current state and future implications. Clin Exp Rheumatol. 2019 Sep-Oct;37 Suppl 120(5):64–72.
4. Dell'Isola A, Allan R, Smith SL, et al. Identification of clinical phenotypes in knee osteoarthritis: a systematic review of the literature. BMC Musculoskelet Disord. 2016 10 12;17(1):425. [PubMed: 27733199]
5. Dell'Isola A, Steultjens M. Classification of patients with knee osteoarthritis in clinical phenotypes: Data from the osteoarthritis initiative. PLoS One. 2018;13(1):e0191045. [PubMed: 29329325]
6. Mobasheri A, van Spil WE, Budd E, et al. Molecular taxonomy of osteoarthritis for patient stratification, disease management and drug development: biochemical markers associated with emerging clinical phenotypes and molecular endotypes. Curr Opin Rheumatol. 2019 1;31(1):80–89. [PubMed: 30461544]
7. van Spil WE, Bierma-Zeinstra SMA, Deveza LA, et al. A consensus-based framework for conducting and reporting osteoarthritis phenotype research. Arthritis Res Ther. 2020 3 20;22(1):54. [PubMed: 32192519] * An expert consensus-based framework for the conduct and reporting of phenotype research in OA
8. Tiulpin A, Thevenot J, Rahtu E, et al. Automatic Knee Osteoarthritis Diagnosis from Plain Radiographs: A Deep Learning-Based Approach. Sci Rep. 2018 1 29;8(1):1727. [PubMed: 29379060] * An approach to automate radiographic diagnosis of knee OA using deep learning
9. Guan B, Liu F, Haj-Mirzaian A, et al. Deep learning risk assessment models for predicting progression of radiographic medial joint space loss over a 48-MONTH follow-up period. Osteoarthritis Cartilage. 2020 4;28(4):428–437. [PubMed: 32035934] ** Deep learning application to predict progression of medial joint space narrowing over 4 years
10. Tiulpin A, Klein S, Bierma-Zeinstra SMA, et al. Multimodal Machine Learning-based Knee Osteoarthritis Progression Prediction from Plain Radiographs and Clinical Data. Sci Rep. 2019 12 27;9(1):20038. [PubMed: 31882803] ** A multi-modal approach using various data types to predict structural progression
11. Liu F, Zhou Z, Samsonov A, et al. Deep Learning Approach for Evaluating Knee MR Images: Achieving High Diagnostic Performance for Cartilage Lesion Detection. Radiology. 2018 10;289(1):160–169. [PubMed: 30063195]
12. Pedoia V, Lee J, Norman B, et al. Diagnosing osteoarthritis from T2 maps using deep learning: an analysis of the entire Osteoarthritis Initiative baseline cohort. Osteoarthritis Cartilage. 2019 7;27(7):1002–1010. [PubMed: 30905742]
13. Tolpadi AA, Lee JJ, Pedoia V, et al. Deep Learning Predicts Total Knee Replacement from Magnetic Resonance Images. Sci Rep. 2020 4 14;10(1):6371. [PubMed: 32286452] ** Use of deep learning on MR images to predict receipt of TKR
14. Nelson AE, Fang F, Arbeeva L, et al. A machine learning approach to knee osteoarthritis phenotyping: data from the FNIH Biomarkers Consortium. Osteoarthritis Cartilage. 2019 7;27(7):994–1001. [PubMed: 31002938] ** Application of novel ML methods to predict pain and structure progression in knee OA
15. Carlesso LC, Neogi T. Identifying pain susceptibility phenotypes in knee osteoarthritis. Clin Exp Rheumatol. 2019 Sep-Oct;37 Suppl 120(5):96–99.** Assessment of pain susceptibility phenotypes, demonstrating the key role of quantitative sensory testing
16. Jiang X, Nelson AE, Cleveland RJ, et al. A precision medicine approach to develop and internally validate optimal exercise and weight loss treatments for overweight and obese adults with knee osteoarthritis. Arthritis Care Res (Hoboken). 2020 3 7;In press.