LEADING ARTICLE

Advancing Rheumatology Care Through Machine Learning

Thomas Hügle[1]

Abstract

Rheumatologic diseases are marked by their complexity, involving immune-, metabolic- and mechanically mediated processes which can affect different organ systems. Despite a growing arsenal of targeted medications, many rheumatology patients fail to achieve full remission. Assessing disease activity remains challenging, as patients prioritize different symptoms and disease phenotypes vary. This is also reflected in clinical trials where the efficacy of drugs is not necessarily measured in an optimal way with the traditional outcome assessment. The recent COVID-19 pandemic has catalyzed a digital transformation in healthcare, embracing telemonitoring and patient-reported data via apps and wearables. As a further driver of digital medicine, electronic medical record (EMR) providers are actively engaged in developing algorithms for clinical decision support, heralding a shift towards patient-centered, decentralized care. Machine learning algorithms have emerged as valuable tools for handling the increasing volume of patient data, promising to enhance treatment quality and patient well-being. Convolutional neural networks (CNN) are particularly promising for radiological image analysis, aiding in the detection of specific lesions such as erosions, sacroiliitis, or osteoarthritis, with several FDA-approved applications. Clinical predictions, including numerical disease activity forecasts and medication choices, offer the potential to optimize treatment strategies. Numeric predictions can be integrated into clinical workflows, allowing for shared decision making with patients. Clustering patients based on disease characteristics provides a personalized care approach. Digital biomarkers, such as patient-reported outcomes and wearables data, offer insights into disease progression and therapy response more flexibly and outside patient consultations. In association with patient-reported outcomes, disease-specific digital biomarkers via image recognition or single-camera motion capture enables more efficient remote patient monitoring. Digital biomarkers may also play a major role in clinical trials in the future as continuous, disease-specific outcome measurement facilitating decentralized studies. Prediction models can help with patient selection in clinical trials, such as by predicting high disease activity. Efforts are underway to integrate these advancements into clinical workflows using digital pathways and remote patient monitoring platforms. In summary, machine learning, digital biomarkers, and advanced imaging technologies hold immense promise for enhancing clinical decision support and clinical trials in rheumatology. Effective integration will require a multidisciplinary approach and continued validation through prospective studies.

1 Introduction

1.1 Clinical Needs in Rheumatology

Rheumatologic diseases are notably heterogeneous due to underlying immune-mediated metabolic or mechanical processes. This can lead to joint destruction such as in rheumatoid arthritis (RA), or can affect various organs, as seen in systemic lupus erythematosus (SLE) and other collagen vascular diseases. Additional factors, such as depression, fibromyalgia, calcifications, or osteoarthritis, often play a role, adding to pain and inflammation stimuli [1]. Diverse patho-mechanisms operate at the cellular level, and even within the same disease, they can vary [2]. For example, RA can exhibit a lymphocytic proliferation on a T- and B-cell level with an invasive reaction of fibroblasts, or alternatively, pauci-cellular macrocytic inflammatory reactions or fibrosis [3]. In return, each pathotype reacts differently to targeted disease-modifying therapies [4]. This diversity is reflected at the cytokine level, with varying responses to cytokine blockade (e.g., tumor necrosis factor [TNF], interleukin [IL]-6,

✉ Thomas Hügle
  Thomas.Hugle@chuv.ch

1 Department of Rheumatology, University Hospital Lausanne (CHUV) and University of Lausanne, Avenue Pierre-Decker 4, 1001 Lausanne, Switzerland

Key Points

Automated image recognition by convolutional neural networks is already used in clinical practice in radiology.

Digital biomarkers and patient-reported outcomes are cornerstones of telemedicine and strengthen flare prediction.

Electronic medical records are a major source of clinical data that can be used to train machine learning models.

Disease prediction of clinical disease course and clustering foster personalized treatment in rheumatic diseases.

IL-17, or IL-23) or cell depletion (CD20) [5]. Despite an ever-expanding arsenal of medications, in most clinical trials, two-thirds of rheumatology patients still do not achieve complete remission [6]. It is crucial to note that assessing disease activity and defining remission in rheumatology is not straightforward. RA alone has a plethora of indicators, such as DAS28, ACR50, Clinical disease activity index (CDAI), EULAR/ACR remission, etc. [7]. Additionally, patients have diverse priorities, with improvements in fatigue or morning stiffness being major concerns, although their measurement remains subjective. In this sense, we mostly follow the treatment recommendations of EULAR/ACR, but this is a systematic, but so far not at all personalized, approach to treating RA or other rheumatic diseases [8]. Meeting these needs requires time, which is unfortunately insufficient due to a shortage of rheumatologists [9]. Other clinical needs include better interdisciplinary coordination of treatment, for example for psoriatic arthritis or arthritis in chronic inflammatory bowel disease. For various reasons, we still think and act in silos, which does not necessarily benefit patients and health care professionals. Finally, there are several rheumatic diseases, such as osteoarthritis of the hand or fibromyalgia, for which no satisfactory drug treatment is yet available [10].

1.2 Digital Transformation

The COVID-19 pandemic has ushered in the era of telemonitoring. In the United States, for instance, telemonitoring of patient-reported outcomes is financially rewarded as a quality assurance measure [11]. Increasingly, patients collect data through apps or wearables, some of which are compatible with electronic medical records (EMRs) and are stored there [12]. This integration of structured and unstructured clinical data with radiological, laboratory, and immunological data helps create a more comprehensive and personalized profile of rheumatology patients, facilitating better prognosis [13]. EMR providers have recognized this trend and are increasingly involved in developing algorithms that can be

integrated directly into EMRs as clinical decision support tools [14]. All of these advancements lay the foundation for a more patient-centered and decentralized approach to medicine, addressing the shortage of rheumatologists or radiologists. Alongside personalization, automation is a fundamental trend, allowing the automation of simple processes such as prescription refills for stable cases or dynamic scheduling through digital tools. Communication with patients can also involve chatbots, deep learning algorithms or advanced language models like ChatGPT4 to handle certain needs, and potentially appointment rescheduling or insurance inquiries [15]. Finally, the internet of things (IOT) and improving sensor and camera technology permit connected medicine with wearables or motion capturing and measurement of mobility as functional biomarkers [16].

1.3 Machine Learning

We are amassing an overwhelming amount of data from our patients, beyond human capacity to manage. This is evident in the expanding array of digital clinical outcome assessments, digital biomarkers, or -omics data, which are now being collected through self-sampling from home along with different imaging modalities that increasingly include photos and videos by patients and doctors [17]. Machine learning allows us to build models to learn from previous data in order to deliver predictions or image recognition [18]. Most of the time, supervised learning models are applied, meaning that the model obtains labeled data (e.g. x-rays with information on erosions, osteophytes etc.). After definition of the output variable of the algorithm, datasets are then divided into a training set (typically 80%) and a test set (20%). The output variable usually is defined as a classification task, such as remission at next visit or a radiographic finding yes/no. In the field of osteoarthritis, this is the case in 90% of the currently published studies (Fig. 1) [19]. Each model should be validated in an independent dataset that is representative for the population where it is applied. Unsupervised learning, that is, algorithms based on unlabeled data, is less frequently applied than supervised learning. Unsupervised learning is used for clustering such as defining disease phenotypes or finding patient outliers in EMRs [20]. Finally, reinforcement learning is the third pillar of machine learning. Reinforcement learning is based on a reward function, or in other words on trial and error. Here, the algorithm is allowed to make new decisions, but has to learn from mistakes (and successful decisions) [21]. Quite advanced in diabetes mellitus, reinforcement learning is not yet applied in the field of rheumatology [21]. In diabetes, it can be applied in a closed system with a simple biomarker (blood glucose) and a simple intervention (insulin injection). A scenario where reinforcement learning could be applied in rheumatology could be in relatively simple tasks,

such as adjusting cortisone doses under strict rules in terms of dose, frequency, etc. Based on the clinical or laboratory response, the model, here also called 'agent', would perform or propose an action that then again will be evaluated by the reward function.

Transfer learning is another machine learning method where a model already developed for a task is reused in another task. Transfer learning is a popular approach in deep learning, as it enables the training of deep neural networks with less data compared with having to create a model from scratch.

In general, the more data that is available to train a machine learning model, the better. However, it's important to recognize that more data doesn't necessarily equate to better data. Clinical judgment and data pre-selection remain crucial, and we must correctly assess data quality and validate algorithm applicability across different patient groups. Currently, over 500 algorithms are FDA approved, with the majority focused on imaging in radiology, cardiology and pathology [22]. Table 1 overviews the different domains of AI.
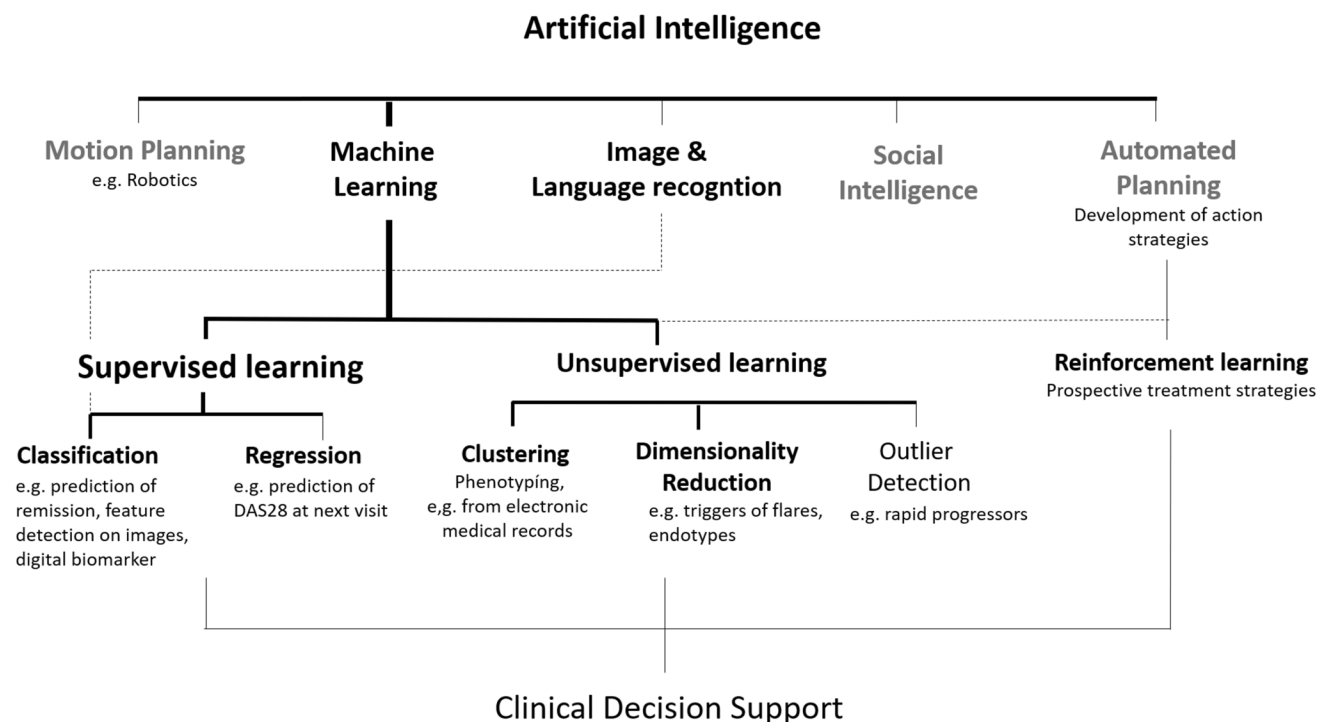
This article begins by discussing imaging, where AI has made significant progress in clinical applications, and then delve into clinical prediction and digital bio-markers. Finally, we will discuss the integration of these applications into the clinical workflow and in clinical trials.

2 Image Recognition

2.1 Clinical Context

In rheumatology, we regularly perform radiographs to detect long-term damage of arthritis, we use ultrasound to assess inflammation during almost every consultation, and typically employ magnetic resonance imaging (MRI) to evaluate spinal structures or sacroiliac joints in spondylarthritis. Clinicians often rely on the radiology department's assessments, but radiology reports are not always promptly available, and younger radiologists may be less familiar with rheumatologic pathologies. Even for rheumatologists, due to effective treatment options, erosions are not always easy to detect. Another common challenge is identifying sacroiliitis in axial spondylarthritis. It has become increasingly clear that we may have previously over-diagnosed mechanical bone marrow edema as spondylarthritis [23]. Automatic computer support based on

Artificial Intelligence

Motion Planning
e.g. Robotics

Machine Learning

Image & Language recognition

Social Intelligence

Automated Planning
Development of action strategies

Supervised learning

Unsupervised learning

Reinforcement learning
Prospective treatment strategies

Classification
e.g. prediction of remission, feature detection on images, digital biomarker

Regression
e.g. prediction of DAS28 at next visit

Clustering
Phenotyping, e,g. from electronic medical records

Dimensionality Reduction
e.g. triggers of flares, endotypes

Outlier Detection
e.g. rapid progressors

Clinical Decision Support

Fig. 1 Overview of domains in artificial intelligence. Supervised machine learning is by far the most widely used, mostly through labelled clinical data such as X-ray images. This is followed by unsupervised learning with unlabeled data, e.g. electronic medical records. Reinforcement learning allows algorithms to make their own decisions and correct them. Complex language models such as Chat-GPT use a combination of supervised, unsupervised, and reinforcement learning

**Table 1** Domains of artificial intelligence (AI)

Domain	Description
Algorithm	A set of step-by-step instructions designed to perform a specific task or solve a particular problem
Machine learning	Focuses on algorithms that learn and make predictions on data (learning from previous experiences)
Natural language processing	Enables machines to understand, interpret, and generate human language
Computer vision	Enabling computers to interpret and process visual data, mimicking the capabilities of human vision. Includes tasks like image and video recognition, object detection, and scene reconstruction
Speech recognition	Technology that recognizes and processes human speech
Planning, scheduling, and optimization	Deals with creating systems that can plan or schedule tasks and optimize processes
Reinforcement learning	Learning method that rewards desired behaviors and/or punishes undesired ones
Neural networks	A subset of machine learning involving models that mimic the human brain's structure and function
Deep learning	A subset of machine learning involving algorithms inspired by the structure and function of the brain's neural networks
Supervised learning	Learning method where the algorithm is trained on labeled data
Unsupervised learning	Learning method where the algorithm learns from unlabeled data to find hidden patterns
Transfer learning	A machine learning technique where a model developed for one task is reused as the starting point for a model on a second, related task
Digital biomarkers	Use of digitally collected data (computer vision or machine learning) that correlate with a disease status
Generative AI	Refers to the type of AI that can generate new content, ideas, or data that are novel and realistic, based on its training data
Vision transformers	Type of neural network model that applies the transformer architecture, originally designed for natural language processing tasks, to analyze and process visual data like images
Radiomics	Extraction of a large number of quantitative features from medical imaging data for diagnostic, prognostic, or predictive purposes

expert opinions would be invaluable in such cases, provided it doesn't consume excessive time [24].

2.2 Convolutional Neural Networks (CNN)

CNN are the primary AI tools for analyzing image data. CNN algorithms analyze radiographs, MRIs, but also photos by assigning a specific task, usually classification, such as detecting erosions, sacroiliitis, or grading arthritis severity [24]. The process begins by scanning images with kernels, which are small quadrants that search for specific features like slanted lines, straight lines, circles, etc. This process is called feature extraction. These 'meta' images with identified features are referred to as convolutions (Fig. 2). Kernels can be two-dimensional or three-dimensional, denoted as 2D or 3D. Convolutional layers are further simplified by kernels into pooling layers. Finally, these images, reduced to specific features, are sent into a neural network responsible for classification. The algorithm's learning process is performed using training and test sets as mentioned above. Images can be automatically optimized before calculation (including data augmentation), and the algorithm may be provided information on regions of interest (segmentation) if necessary.

Classification quality is typically measured through algorithm accuracy or, with sufficient data, sensitivity and specificity. In rheumatology, numerous CNN algorithms exist, with some being FDA approved, such as those for detecting and scoring knee osteoarthritis or spondylarthritis [25]. No-coding platforms now allow users to upload, segment, and classify images without coding knowledge, even creating web apps to directly apply the generated algorithm [26]. Algorithm improvement can be achieved through preprocessing, highlighting specific features in images, such as finger creases for swelling detection or hip contours in X-rays. Some algorithms integrate clinical data with radiological images to predict radiological outcomes [27]. The results are sometimes poor in terms of accuracy (<60%), raising questions about the practicality of using MRIs for such purposes or whether patients would receive joint replacements or alternative therapies based on these results. Transfer learning can be used to leverage CNN models by using the knowledge acquired from a previously learned similar task [28]. This approach has significantly impacted medical image analysis by addressing the challenges of limited data availability and reducing the need for extensive time and computational resources.
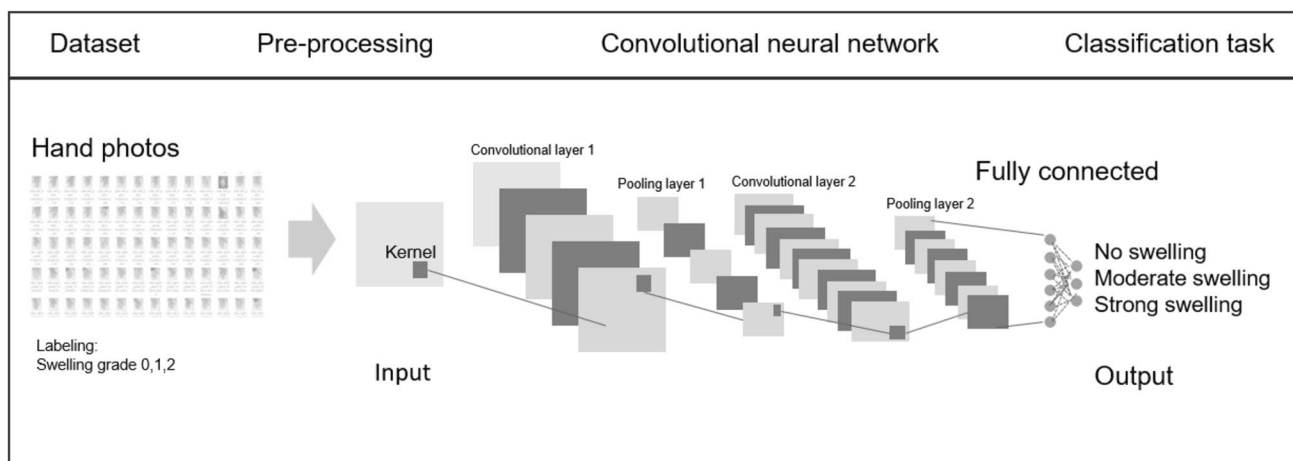
Fig. 2 Convolutional neural network. Here with a classification task to predict joint swelling from hand photos of patients with rheumatoid arthritis

2.3 Auto-Machine Learning Platforms

To train CNN algorithms, automated machine learning platforms, also called autoML or no-coding platforms, can now be used [26]. Here you drag and drop the images with the corresponding labels onto the platform and select a CNN architecture, such as Resnet34. The platform automatically augments the images beforehand to increase the performance of the algorithm. Later, a user interface in the form of a web app can be downloaded directly from the platform. This makes it possible for clinicians and scientists who do not code to create algorithms and test them, for example, for usefulness or usability [29]. This allows, for example, models for niche tasks to be trained on small data sets. Or, in preclinical studies, images can be evaluated more quickly than by hand.

2.4 Vision Transformers

Vision Transformers (ViT) are a newer type of neural network model designed for image recognition tasks, inspired by the success of transformers in NLPs such as ChatGPT (where the letter T stands for 'Transformer') [30]. Unlike traditional CNNs that process images using local features, ViTs divide an image into fixed-size patches and flatten these patches into a sequence, similar to words in a sentence. Each patch is then encoded with positional information. The transformer architecture, with its self-attention mechanism, processes these sequences, allowing the model to weigh the importance of different patches (small segments or portions of an input data set) in relation to each other. The primary disadvantage of ViTs compared with CNNs is their requirement for a larger amount of data to achieve optimal performance. Unlike CNNs, which have inductive biases such as translation invariance and locality that make them naturally suited for image data and allow them to perform well even with relatively less data, ViTs lack these biases. As a result, they need substantial training data to learn these features implicitly. Furthermore, ViTs are generally more computationally intensive and require more resources for training. This is due to the self-attention mechanism in transformers, which scales quadratically with the number of image patches, leading to higher memory usage and longer training times, especially for large images.

2.5 Radiomics

Radiomics is an emerging field in medical imaging that involves the conversion of images into high-dimensional, quantifiable data [31]. This process is achieved through the extraction of a large number of features from medical imaging scans such as computed tomography (CT), MRI, and positron emission tomography (PET). These features, which are not readily apparent to the human eye, include details about the shape, texture, intensity, and the overall architecture of the image. Radiomics aims to uncover patterns within this data that are relevant for disease diagnosis, prognosis, and predicting treatment response. The role of AI, particularly machine learning, in radiomics is pivotal. AI algorithms are adept at handling and interpreting the vast and complex data generated in radiomics. They can efficiently process these high-dimensional datasets to identify subtle patterns and correlations that are beyond human analytical capability. For instance, in oncology, radiomic features extracted from tumor images can be analyzed using AI to differentiate between benign and malignant tumors, determine the tumor stage, and predict the patient's response to certain therapies. As an example, in rheumatology, radiomic analysis of high

resolution computed tomography (HRCT) has shown to predict mortality in RA patients with interstitial lung disease and may promote HRCT as a digital biomarker [32].

3 Clinical Predictions

3.1 Clinical Context

Rheumatic diseases like RA often exhibit fluctuating and challenging-to-assess clinical courses. Drug survival can be short, requiring rheumatologists to constantly adapt treatments. International treatment guidelines, such as EULAR criteria, advocate for a 'Treat-to-Target' strategy, aiming to achieve low disease activity within 3–6 months [8]. However, this approach can lead to delays, with patients remaining on a medication for an extended period before it is discontinued. Precise disease prognosis, or ideally, selecting the right medication, is desirable.

3.2 Disease Prediction and Clustering

Machine learning is a powerful and flexible tool for clinical predictions. Predicting disease activity in the form of a numeric value, such as Disease Activity Score-28 for Rheumatoid Arthritis with C-Reactive Protein (DAS28-CRP), involves a regression analysis. We have employed a new deep learning architecture in a dataset involving approximately 12,000 Swiss RA patients and predicted DAS28-CRP at next visit [33]. An 8% accuracy rate compared with actual values was achieved. Previous studies in the USA, utilizing EMR data, focused on classifications, such as predicting active disease or inactive disease [34]. Predicting numerical values may better integrate into clinical rationale and workflow, allowing for the creation of a dashed line representing disease activity over time. This could be discussed with the patient as part of shared decision making on whether to continue or change treatment, potentially surpassing the 'Treat-to-Target' strategy. Numeric prediction could also be applied to laboratory values like CRP or anti-dsDNA. For instance, a trend arrow on laboratory reports could indicate changes. However, predicting medication choice seems more distant, likely due to a lack of head-to-head studies and convincing evidence for such algorithms. A recent study performed a relatively straightforward prediction: non-response to methotrexate [35]. Since methotrexate is the common first-line therapy, these data could be more effectively utilized. A high likelihood of non-response might qualify a patient for biologic or targeted synthetic disease-modifying antirheumatic drug (DMARD) treatment directly.

Clustering involves unsupervised learning using unlabeled data to form clusters or phenotypes that appear similar. Our own study demonstrated that drug survival for tocilizumab and clinical response differs between clusters. It's important to note that patients may change clusters during their patient journey, making predictions more challenging. There may not be a dedicated 'tocilizumab cluster', for example [36]. Clinical, biological, and radiological data can also be combined to identify disease 'endotypes'. In a compelling study of osteoarthritis, unsupervised learning techniques were applied for clustering analysis [37]. The study involved pooling and preprocessing clinical data from questionnaires and imaging, as well as biochemical information from blood and urine samples. A model was then trained using principal component analysis and k-means clustering. The findings were subsequently corroborated using traditional statistical methods, including the Mann-Whitney U test and chi-square test.

3.3 Digital Biomarkers

To best predict clinical outcomes, including treatment choices, disease-specific data along the patient journey are essential [38]. Currently, digital biomarkers primarily encompass patient-reported outcomes (PROs) or data from wearables, either stored by patients or ideally transmitted to EMRs through application programming interfaces (APIs) [39]. These data are valuable but not necessarily disease-specific. For instance, pain or fatigue is highly sensitive to influences like depression or fibromyalgia [40]. Home blood sampling may enhance specificity, but it involves organizational challenges and remains invasive [17]. Other sensor technologies will likely emerge, such as non-invasive CRP determination or thermal cameras.

User-friendly telemonitoring technologies are also available for patients, such as utilizing smartphone cameras [41]. In our own study, DETECTRA involves the automatic recognition of finger creases as biomarkers for joint swelling [42]. Changes in finger creases occur during synovitis or periarticular swelling, creating an 'inflammatory fingerprint'. This process is reversible, allowing us to detect not only inflammatory flares but also treatment responses. The process involves three steps: detecting the hand in a photo using keypoint detection, isolating the desired joints (e.g. finger joints), and recognizing and measuring finger folds (Fig. 3). This can be achieved through computer vision techniques such as Canny Edge Detection or Ridge Detection, although we achieved better results with a newly trained CNN considering crease pixel length and diameter. Another project uses single-camera motion detection to assess finger joint mobility. Finger joint movements are captured on a normal mobile phone camera and angles and speed are measured and (theoretically) transferred to the EMR (Fig. 4) [43].
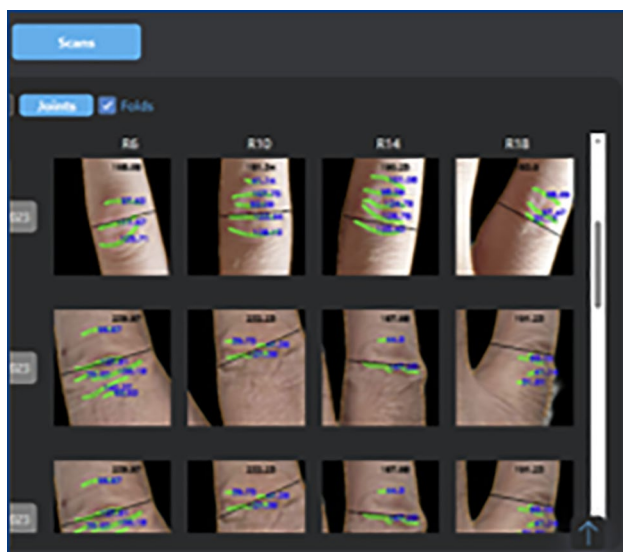
Fig. 3 A dashboard for remote telemonitoring showing automated measurement of finger folds and joint diameters over time to detect arthritis flares. Images were taken by patients at home on their mobile phone app in combination with information on joint pain and stiffness

4 Integration of AI in the Clinical Workflow

4.1 Imaging

Image recognition is probably the easiest machine learning modality to be included in the clinical workflow [44]. This can start with an automatically generated radiology report. Several algorithms are already FDA-approved and in practice (e.g. detection of fractures or osteoarthritis) [25]. Heatmaps are, in my view, a promising candidate to be integrated into clinical practice. Heatmaps illustrate the region of interested for the algorithm, usually by color, and can guide the clinician to certain lesions [45]. A good example is the classification of osteoarthritis grades (Fig. 4), showing a conserved joint space in grade 0 versus osteophytes in grade 1 or diffuse subchondral bone remodeling in grade 4 [29]. Heatmaps are particularly popular in pathology for the evaluation of histological slides.

4.2 Digital Pathways and Remote Monitoring

The clinical workflow for rheumatic diseases involves processes from diagnosis to therapy and monitoring [43]. Given the chronic nature of many rheumatic diseases, a cycle emerges where either no treatment modification occurs, or treatment is adjusted. Diagnostic processes include routine examinations (e.g., blood tests every 6 months, yearly X-rays) and exceptional investigations (e.g., ultrasound and blood tests during disease flares). Additionally, there are processes for comorbidity assessment, prophylaxis, and
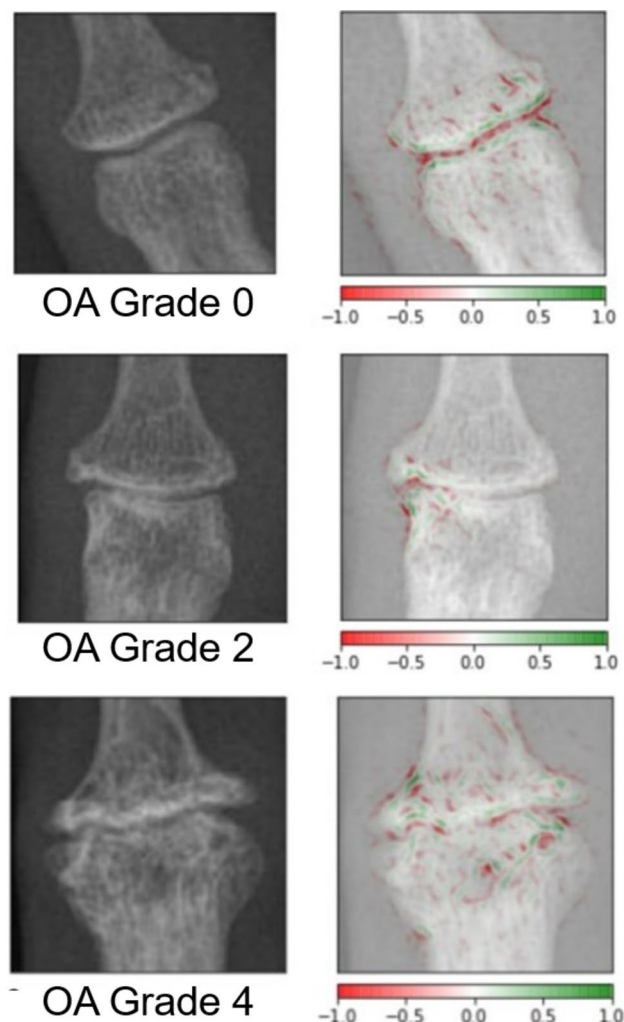
Fig. 4 Heatmaps of a convolutional neural network (CNN) algorithm to classify osteoarthritis (OA) grade in hand osteoarthritis

vaccination. Together, these processes create the patient journey, which can be digitally tracked over time. The existence of digital pathways allows for quality and efficiency assessments using key performance indicators (KPIs) and quality markers. In our case, we use the SCQM Register, which records disease characteristics (currently for rheumatoid arthritis, psoriasis arthritis, spondylarthritis and giant cell arteritis) during doctor visits and through an app. Notwithstanding, remote patient monitoring harbors several challenges. Integrating register data into EMRs via APIs can be difficult both on a technical and regulatory level. In the United States, health insurance covers remote patient telemonitoring services to some extent, a practice that is not as common in many other countries [11]. It is also not yet clear how the Remote Patient Monitoring (RPM) will be organized in practice. In some cases, this is done by health care coaches or nurses, who also call patients at certain intervals.

4.3 Large Language Models (LLM)

Large language models (LLMs) like ChatGPT are currently deployed across industries and also can be invaluable assets in medical care [46]. For administrative tasks, they can streamline paperwork, extract key data from patient records, or generate comprehensive reports. Furthermore, LLMs can facilitate patient education, translating complex medical jargon into layman's terms and answering common queries. This not only empowers patients with knowledge about their conditions but also reduces the burden on medical staff. Moreover, with their multilingual capabilities, LLMs can bridge language barriers, ensuring clear communication between patients and providers. If and how LLMs can serve as decision-support tools in offering diagnostic suggestions or clarifying medical concepts remains to be evaluated, but so far ChatGPT is not a medical device.

5 Clinical Trials

Digital tools are crucial for the success of decentralized clinical trials (DCTs) [47]. DCTs shift from traditional site-centric trials to more patient-centered ones, offering greater flexibility and convenience. With the support of wearables, smartphone apps, and telemedicine platforms, patients can participate in trials without frequent site visits, making it especially advantageous for those with mobility challenges or residing in distant areas. These tools enable real-time remote monitoring, virtual consultations, and direct-to-patient shipments of investigational treatments. Electronic consent platforms streamline patient enrollment, while reminder apps and interactive platforms enhance patient engagement and retention. Furthermore, DCTs, powered by digital tools, can broaden patient enrollment, resulting in more diverse and representative trials [48]. By consolidating data from various sources, digital platforms provide researchers with a holistic view of patient data. Machine learning algorithms are also used for patient selection. For example, patients with a high predicted disease activity are more likely to respond to a drug, and certain disease clusters respond better to treatment than others. Machine learning algorithms can also be used to reduce sample sizes, improve enrollment, and conduct faster, more optimized adaptive clinical trials [49].

6 Drug Development

Drug development is not directly involved in clinical decision making, but nevertheless represents an important area for AI in medicine; AI may reduce the time and cost of bringing new drugs to market [50]. AI algorithms rapidly sift through vast datasets to identify promising drug candidates and suggest novel compounds. It also plays an increasing role in target identification by analyzing biological data to uncover and validate new drug targets, essential for diseases with complex pathology [51]. Machine learning algorithms can speed up the compound screening process, analyzing thousands of compounds quickly to determine their efficacy against specific targets, reducing the time and costs compared with traditional methods. Additionally, AI models are adept at predicting the efficacy and safety profiles of drug candidates before clinical trials, reducing the risks of failure in later stages [52].

7 Ethical Aspects

AI algorithms solve very specific tasks. How relevant and therefore ethical these tasks are must be assessed individually by doctors, ethics committees and, ideally, patients. The explainability of AI is important. Preprocessing can be used to focus on image recognition, for example, and make algorithms more transparent or relativize the black box. The solutions provided by machine learning algorithms always depend on the quality of the data, the so-called ground truth. It should be ensured that the trained data correctly represents the target population and that no population segments are disadvantaged. Independent clinical studies must test algorithms in independent patient populations.

8 Outlook

Machine learning algorithms are about to enter rheumatology in the form of image recognition, disease prediction, clinical workflow, and clinical trials. The prediction of efficacy and safety of individual drugs remains a challenge, but first steps in the form of cluster response to treatment classes or methotrexate non-response have been taken. The EMR as a main source of real-world data has been identified for AI predictions that are easy to implement in the clinical workflow. Disease-specific digital biomarkers will help make the patient journey more transparent and predictable. Remote telemonitoring will leverage patient care and empowerment in rheumatology but needs substantial reorganization of processes and staff. Generative AI, notably via LLMs, may support the administrative burden.

Declarations

References

1. Katz J, Bartels CM. Multimorbidity in rheumatoid arthritis: literature review and future directions, (in eng). Curr Rheumatol Rep. 2024;26(1):24–35. https://doi.org/10.1007/s11926-023-01121-w.

2. Weisenfeld D et al. Associations between rheumatoid arthritis clinical factors with synovial cell types and states, (in eng), Arthritis Rheumatol, 2023, https://doi.org/10.1002/art.42726.

3. Müller-Ladner U, et al. Synovial fibroblasts of patients with rheumatoid arthritis attach to and invade normal human cartilage when engrafted into SCID mice, (in eng). Am J Pathol. 1996;149(5):1607–15.

4. Rivellese F, et al. Rituximab versus tocilizumab in rheumatoid arthritis: synovial biopsy-based biomarker analysis of the phase 4 R4RA randomized trial, (in eng). Nat Med. 2022;28(6):1256–68. https://doi.org/10.1038/s41591-022-01789-0.

5. Kondo N, Kuroda T, Kobayashi D (2021) Cytokine networks in the pathogenesis of rheumatoid arthritis. (in eng), Int J Mol Sci, 22(20), https://doi.org/10.3390/ijms222010922.

6. Lauper K, et al. Effectiveness of TNF-inhibitors, abatacept, IL6-inhibitors and JAK-inhibitors in 31 846 patients with rheumatoid arthritis in 19 registers from the "JAK-pot" collaboration, (in eng). Ann Rheum Dis. 2022;81(10):1358–66. https://doi.org/10.1136/annrheumdis-2022-222586.

7. Pappas DA, et al. Treatment patterns and clinical outcomes in patients with rheumatoid arthritis initiating etanercept, adalimumab, or Janus kinase inhibitor as first-line therapy: results from the real-world CorEvitas RA Registry, (in eng). Arthritis Res Ther. 2023;25(1):166. https://doi.org/10.1186/s13075-023-03120-9.

8. Smolen JS, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update, (in eng). Ann Rheum Dis. 2017;76(6):960–77. https://doi.org/10.1136/annrheumdis-2016-210715.

9. Kulhawy-Wibe SC, et al. Results from the 2020 Canadian rheumatology association's workforce and wellness survey, (in eng). J Rheumatol. 2022;49(6):635–43. https://doi.org/10.3899/jrheum.210990.

10. Kloppenburg M, et al. Etanercept in patients with inflammatory hand osteoarthritis (EHOA): a multicentre, randomised, double-blind, placebo-controlled trial, (in eng). Ann Rheum Dis. 2018;77(12):1757–64. https://doi.org/10.1136/annrheumdis-2018-213202.

11. Register F Medicare Program; CY 2020 Revisions to Payment Policies Under the Physician Fee Schedule and Other Changes to Part B Payment Policies; Medicare Shared Savings Program Requirements; Medicaid Promoting Interoperability Program Requirements for Eligible Professionals; Establishment of an Ambulance Data Collection System; Updates to the Quality Payment Program; Medicare Enrollment of Opioid Treatment Programs and Enhancements to Provider Enrollment Regulations Concerning Improper Prescribing and Patient Harm; and Amendments to Physician Self-Referral Law Advisory Opinion Regulations." https://www.federalregister.gov/documents/2019/08/14/2019-16041/medicare-program-cy-2020-revisions-to-payment-policies-under-the-physician-fee-schedule-and-other accessed.

12. Uslu A, Stausberg J. Value of the electronic medical record for hospital care: update from the literature, (in eng). J Med Internet Res. 2021;23(12):e26323. https://doi.org/10.2196/26323.

13. Curtis JR, Ford K, Fiore S, Mikuls TR, Shaney K, Sethi V Development of an Interactive Tool to Support Shared Decision Making in Rheumatoid Arthritis: Treatment Attribute Preference, (in eng), *Arthritis Care Res (Hoboken),* 2023, doi: https://doi.org/10.1002/acr.25223.

14. Cosmos. https://cosmos.epic.com/ accessed.

15. Entrepreneur NC AI aims to improve access to NHS appointments. https://nhscep.com/2023/11/14/ai-aims-to-improve-access-to-nhs-appointments/ accessed.

16. Latif S et al. Deep Learning for the Industrial Internet of Things (IIoT): A Comprehensive Survey of Techniques, Implementation Frameworks, Potential Applications, and Future Directions, (in eng), *Sensors (Basel)* 21(22) 2021 https://doi.org/10.3390/s21227518.

17. Muehlensiepen F, et al. At-home blood self-sampling in rheumatology: a qualitative study with patients and health care professionals, (in eng). BMC Health Serv Res. 2022;22(1):1470. https://doi.org/10.1186/s12913-022-08787-5.

18. Hügle M, Omoumi P, van Laar JM, Boedecker J, Hügle T (2020) Applied machine learning and artificial intelligence in rheumatology, (in eng), Rheumatol Adv Pract 4(1):rkaa005 https://doi.org/10.1093/rap/rkaa005.

19. Binvignat M et al. Use of machine learning in osteoarthritis research: a systematic literature review, (in eng), RMD Open 8(1) 2022, https://doi.org/10.1136/rmdopen-2021-001998.

20. Knevel R, Liao KP. From real-world electronic health record data to real-world results using artificial intelligence, (in eng). Ann Rheum Dis. 2023;82(3):306–11. https://doi.org/10.1136/ard-2022-222626.

21. Hügle T. Learning from chess engines: how reinforcement learning could redefine clinical decision-making in rheumatology, (in eng). Ann Rheum Dis. 2022;81(8):1072–5. https://doi.org/10.1136/annrheumdis-2022-222141.

22. FDA. https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices accessed.

23. Marnas G, et al. Spine abnormalities associated with bone edema on sacroiliac joints MRI in patients with non-inflammatory

chronic back pain, (in eng). Joint Bone Spine. 2022;89(6): 105436. https://doi.org/10.1016/j.jbspin.2022.105436.

24. Lee S, et al. Artificial intelligence for the detection of sacroiliitis on magnetic resonance imaging in patients with axial spondyloarthritis, (in eng). Front Immunol. 2023;14:1278247. https://doi.org/10.3389/fimmu.2023.1278247.

25. FDA. 2023. https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices accessed.

26. Faes L, et al. Automated deep learning design for medical image classification by health-care professionals with no coding experience: a feasibility study, (in eng). Lancet Digit Health. 2019;1(5):e232–42. https://doi.org/10.1016/s2589-7500(19)30108-6.

27. Schiratti JB, et al. A deep learning method for predicting knee osteoarthritis radiographic progression from MRI, (in eng). Arthritis Res Ther. 2021;23(1):262. https://doi.org/10.1186/s13075-021-02634-4.

28. Kora P, et al. Transfer learning techniques for medical image analysis: a review. Biocybernetics Biomed Eng. 2022;42(1):79–107. https://doi.org/10.1016/j.bbe.2021.11.004.

29. Caratsch L et al. POS0892 An End-to-End Machine Learning Pipeline for the Automated Detection of Radiographic Hand Osteoarthritis: A No-coding Platform Experience, vol. 82, ed. Annals of the Rheumatic Diseases, 2023, pp. 753-754.

30. Khan RF, Lee BD, Lee MS. Transformers in medical image segmentation: a narrative review, (in eng). Quant Imaging Med Surg. 2023;13(12):8747–67. https://doi.org/10.21037/qims-23-542.

31. Logotheti S, Georgakilas AG (2023) More than Meets the Eye: Integration of Radiomics with Transcriptomics for Reconstructing the Tumor Microenvironment and Predicting Response to Therapy, in Cancers (Basel), 15(6). Switzerland

32. Venerito V, et al. Radiomics to predict the mortality of patients with rheumatoid arthritis-associated interstitial lung disease: a proof-of-concept study, (in eng). Front Med (Lausanne). 2022;9:1069486. https://doi.org/10.3389/fmed.2022.1069486.

33. Hügle M et al. Personalized Prediction of Disease Activity in Patients with Rheumatoid Arthritis Using an Adaptive Deep Neural Network, ed. PLOS One. In press. 2021

34. Norgeot B et al. Assessment of a deep learning model based on electronic health record data to forecast clinical outcomes in patients with rheumatoid arthritis, (in eng), JAMA Netw Open, 29(3): e190606, 03 2019, https://doi.org/10.1001/jamanetworkopen.2019.0606.

35. Duquesne J, et al. Machine learning identifies a profile of inadequate responder to methotrexate in rheumatoid arthritis, (in eng). Rheumatology (Oxford). 2023;62(7):2402–9. https://doi.org/10.1093/rheumatology/keac645.

36. Kalweit M, Burden AM, Boedecker J, Hügle T, Burkard T. Patient groups in Rheumatoid arthritis identified by deep learning respond differently to biologic or targeted synthetic DMARDs, (in eng). PLoS Comput Biol. 2023;19(6): e1011073. https://doi.org/10.1371/journal.pcbi.1011073.

37. Angelini F, et al. Osteoarthritis endotype discovery via clustering of biochemical marker data, (in eng). Ann Rheum Dis. 2022;81(5):666–75. https://doi.org/10.1136/annrheumdis-2021-221763.

38. Mucke J, Krusche M, Burmester GR A broad look into the future of rheumatoid arthritis, (in eng), Ther Adv Musculoskelet Dis 14: 1759720x221076211, 2022, https://doi.org/10.1177/1759720x221076211.

39. Gossec L, et al. Detection of flares by decrease in physical activity, collected using wearable activity trackers in rheumatoid arthritis or axial spondyloarthritis: an application of machine learning analyses in rheumatology, (in eng). Arthritis Care Res (Hoboken). 2019;71(10):1336–43. https://doi.org/10.1002/acr.23768.

40. Yuan SLK, Marques AP. Development of ProFibro—a mobile application to promote self-care in patients with fibromyalgia, (in eng). Physiotherapy. 2018;104(3):311–7. https://doi.org/10.1016/j.physio.2018.04.005.

41. Kusunose M et al. (2023) Measurement of shoulder abduction angle with posture estimation artificial intelligence model, (in eng), Sensors (Basel), 23(14) https://doi.org/10.3390/s23146445.

42. Hügle T et al. (2022) Dorsal finger fold recognition by convolutional neural networks for the detection and monitoring of joint swelling in patients with rheumatoid arthritis. (in eng), Digit Biomark, 6(2): 31-35 https://doi.org/10.1159/000525061.

43. Hamann PDH, Minaur N, Tobias JH, Clark EM. Capturing remote disease activity—results of a 12-month clinical pilot of a smartphone app in NHS rheumatology clinics in Bristol, (in eng). Rheumatology (Oxford). 2020;59(8):2158–61. https://doi.org/10.1093/rheumatology/keaa015.

44. Mohsen F, Ali H, El Hajj N, Shah Z. Artificial intelligence-based methods for fusion of electronic health records and imaging data, (in eng). Sci Rep. 2022;12(1):17981. https://doi.org/10.1038/s41598-022-22514-4.

45. Hong S et al. Methods of visualizing the results of an artificial-intelligence-based computer-aided detection system for chest radiographs: effect on the diagnostic performance of radiologists, (in eng), Diagnostics (Basel) 13(6) 2023, https://doi.org/10.3390/diagnostics13061089.

46. Hügle T The wide range of opportunities for large language models such as ChatGPT in rheumatology, (in eng), RMD Open 9(2), 2023, https://doi.org/10.1136/rmdopen-2023-003105.

47. de Jong AJ et al. Opportunities and Challenges for Decentralized Clinical Trial Approaches: European Health Technology Assessment Perspective. (in eng), Value Health, 2023, https://doi.org/10.1016/j.jval.2023.11.006.

48. Sehrawat O, Noseworthy PA, Siontis KC, Haddad TC, Halamka JD, Liu H. Data-driven and technology-enabled trial innovations toward decentralization of clinical trials: opportunities and considerations, (in eng). Mayo Clin Proc. 2023;98(9):1404–21. https://doi.org/10.1016/j.mayocp.2023.02.003.

49. Askin S, Burkhalter D, Calado G, El Dakrouni S. Artificial Intelligence Applied to clinical trials: opportunities and challenges, (in eng). Health Technol (Berl). 2023;13(2):203–13. https://doi.org/10.1007/s12553-023-00738-2.

50. Qureshi R, et al. AI in drug discovery and its clinical relevance, (in eng). Heliyon. 2023;9(7): e17575. https://doi.org/10.1016/j.heliyon.2023.e17575.

51. Pun FW, Ozerov IV, Zhavoronkov A AI-powered therapeutic target discovery. Trends Pharmacol Sci 44(9): 561-572, 2023/09/01/ 2023, https://doi.org/10.1016/j.tips.2023.06.010.

52. Smith GF. Artificial intelligence in drug safety and metabolism, (in eng). Methods Mol Biol. 2022;2390:483–501. https://doi.org/10.1007/978-1-0716-1787-8_22.