



# How Health Information Technologies and Artificial Intelligence May Help Rheumatologists in Routine Practice

Nathan Foulquier · Pascal Redou · Alain Saraux

Received: April 9, 2019 / Published online: April 26, 2019  
© The Author(s) 2019

**Keywords:** Artificial intelligence; Bioinformatics; Machine learning; Rheumatologists

Over the last decades, technologies of signal acquisition in biology have evolved to produce an increasing amount of data. Approaches such as next-generation sequencing, flow cell cytometry and quantitative PCR have led to an explosive increase in the number of possible variables used to describe a patient. Bioinformatics has played a key role in increasing the quantity and quality of available data through

**Enhanced Digital Features** To view enhanced digital features for this article go to: <https://doi.org/10.6084/m9.figshare.7977143>.

N. Foulquier · P. Redou  
Laboratoire de Traitement de l'Information Médicale (LATIM), UMR 1101, Brest Institute of Biological Research (IBRBS), Inserm, Université de Brest–Centre Hospitalier Universitaire, Brest, France

A. Saraux (✉)  
Rheumatology Unit, Centre National de Référence des Maladies Auto-Immunes Rares (CERAINO), Université de Brest–Centre Hospitalier Universitaire, Brest, France  
e-mail: [alain.saraux@chu-brest.fr](mailto:alain.saraux@chu-brest.fr)

A. Saraux  
UMR 1227, Lymphocytes B et Autoimmunité, Inserm, Université de Brest–Centre Hospitalier Universitaire, Brest, France

A. Saraux  
LabEx IGO, Brest, France

applications such as genomic annotation and protein structure prediction. All of these processes contribute to the generation of an ever-increasing amount of data, which in turn will provide the clinician with a better representation of the patient. The introduction of artificial intelligence (AI) approaches (especially machine learning techniques) in recent years provides the field of bioinformatics with the foundation for a real game changer: computer aids are no longer restrained to increment and clean data, which provides the clinician with an even better representation of the patient, but they are now capable of extracting—very efficiently—complex information from these representations, leading to the possibility of offering yet more advanced support tools for various complex tasks. In this commentary we focus on the field of rheumatology and present a list of applications for which there is the potential for prominent use of the next generation of bioinformatics tools.

Diagnostic aids are probably the first category of tools that come to mind when talking about AI in a biomedical context, but creating clinical decision support systems (CDSSs) is a challenging task. For example, classic deep-learning approaches can identify complex relations between variables and perform very well in terms of identifying biological signatures strongly linked to diseases based on clinical and biological data [1], but they often lack

transparency and behave as black boxes, which is unsuitable for CDSSs [2]. Variants of deep-learning architectures using saliency grid computation could be used to “break the black box” [2], but other algorithms, such as decision trees, could also be used to directly suggest a diagnosis and highlight the rules followed to compute the prediction, assuring transparency of the decision-making process. AI techniques can also be used to perform natural language processing (NLP) through the use of recurrent neural networks or even particular types of very deep convolutional network architecture [3]. There are many applications of NLP in the biomedical environment, from the development of vocal interfaces for virtual assistants to the analysis of electronic health records (EHRs) [4]. EHRs represent a very large collection of medical documents containing information about patients; analyzing these documents through NLP is one way to generate high-level data and ameliorate patient representation in the data, which is especially valuable for machine learning techniques.

Systematic literature review is a powerful tool to investigate specific topics. The amount of available information is constantly growing due to the extensively increasing number of articles published each year. However, going through the process of retrieving, evaluating and selecting an important number of articles with the aim to identify and extract the most relevant information can be very time consuming and require meticulous and repeated tasks. The development of NLP techniques and the possibility to easily access publication databases over the internet create the possibility of automatic bibliography screening approaches to perform these tasks with an effectiveness comparable to a manual systematic literature review performed by a clinician [5]. These tools can—within minutes—provide the insights of a scientific consensus over a complex question by analyzing thousands of publications, which is not only a gain of time for the clinician, but can also be seen as a way to retrieve prior knowledge on a specific question, which is a very interesting process in the machine learning context. To date, the immense majority of machine learning techniques rely only on the information

they can find in the datasets provided for the training phase of the algorithm; for example, an algorithm trained to classify patients into two classes will use a collection of labeled patients and will learn during the training a set of rules that separate the two groups most accurately within the provided dataset. Automatic scientific publication processing through NLP techniques could be a way to retrieve pre-existing rules from the literature: instead of learning all of the rules “from scratch,” it would be possible, by using a limited collection of patient data in a dataset, to use pre-existing knowledge of this specific classification problem and take advantage of the decades of scientific expertise present in publications. These kinds of approaches are close to Bayesian learning [6], which could be a great candidate for the next major improvement in AI.

The development of bioinformatics techniques over the last decade has been very important in the development of personal medicine by giving the patient the possibility to generate information on him/her self using disease-specific applications on mobile devices. Such information has a multitude of potential uses: it enables a description of patient evolution over time; it can be used to create a large dataset to be used by machine learning algorithms; it may improve communication between the clinician and the patient; and it also empowers patients in the daily management of their disease to improve patient care. Disease-specific applications for mobile devices could also represent a gain of valuable time for healthcare professionals; for example, rheumatologists could use many applications or internet websites to compute the specific disease activity score [7].

Bioinformatics could also be used to perform standardization of a variety of information, providing clinicians with the possibility to compare information accurately. For example, the use of different activity scores in rheumatology for the same pathology could be standardized with the appropriate tool [8]. The standardized system would then produce a similar wide range of values (e.g. 0–10) and the same cutoffs to define remission and low and high activity (e.g. 1, 3 and 5 for all classification

criteria instead the classical 2.6, 3.2 and 5.1 for the disease activity score in rheumatoid arthritis and 1.5, 7 and 17 for the polymyalgia activity score). Such standardization could facilitate the interpretation of activity scores for various diseases.

Image analysis is the area where machine learning has achieved the most remarkable advances in performance, especially through the use of convolutional neural networks (CNN). From its introduction at the IMAGENET competition in 2012 [9] to the recent release of the NVIDIA Face Generator software (NVIDIA Corp., Santa Clara, CA, USA), this algorithm has found applications in many areas, including rheumatology, with rheumatologists being quite active in utilizing CNN for the analysis of X-rays, ultrasound or magnetic resonance imaging [10]. These advanced imaging skills could also be used to provide a support in technical applications, such as percutaneous needle puncture or biopsy-guided by ultrasound.

Machine learning on large cohorts of treated patients can highlight treatment targets and complex secondary manifestations on unsuspected pathways. If patients' data are available as a time series, the use of residual neural network (ResNET) technology would be appropriate to take advantage of the sequential information. Classical bioinformatic tools could then be used to identify specific treatment resistance in patients by looking at a set of prior identified targets (i.e. specific gene expression or protein concentration) and give an estimation of the effectiveness of a set of treatments.

To conclude, the integration of computer technologies in medicine has been around for a long time in the form of automating the acquisition and processing of patients' digital data. These processes essentially consist of enriching the dataset available to the clinician for the decision-making process regarding a diagnosis. More recently, bioinformatics has taken a new turn with the arrival of AI; the newly developed tools made available to the clinician are no longer there just to provide a better quality of data but also to extract from these data the information needed to characterize the patient's condition, thereby

facilitating the decision-making process towards the correct diagnosis. In the near future we can imagine that in addition to using the information contained in the data at their disposal, diagnostic tools will be able to retrieve external information, in the form of "knowledge" about the pathologies, independent of the data they must process to make a diagnosis.

## ACKNOWLEDGEMENTS

**Funding.** No funding or sponsorship was received for this study or publication of this article.

**Authorship.** All named authors meet the International Committee of Medical Journal Editors (ICMJE) criteria for authorship for this article, take responsibility for the integrity of the work as a whole, and have given their approval for this version to be published.

**Disclosures.** Alain Saraux is a member of the journal's Editorial Board. Nathan Foulquier and Pascal Redou have nothing to disclose.

**Compliance with Ethics Guidelines.** This article is based on previously conducted studies and does not contain any studies with human participants or animals performed by any of the authors.

**Peer Review.** Please note, contrary to the journal's standard single-blind peer review process, as an editorial this article underwent review by a member of the journal's Editorial Board.

**Open Access.** This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

---

## REFERENCES

1. Khanna NN, Jamthikar AD, Gupta D, et al. Rheumatoid arthritis: atherosclerosis imaging and cardiovascular risk assessment using machine and deep learning-based tissue characterization. *Curr Atheroscler Rep.* 2019;21(2):7. <http://link.springer.com/10.1007/s11883-019-0766-x>.
2. Shortliffe EH, Sepúlveda MJ. Clinical decision support in the era of artificial intelligence. *JAMA.* 2018;320(21):2199–200.
3. Conneau A, Schwenk H, Barrault L, Lecun Y (2016) Very deep convolutional networks for natural language processing. 2016;2. arXiv:1606.01781 [cs.CL].
4. Esteva A, Robicquet A, Ramsundar B, et al. A guide to deep learning in healthcare. *Nat Med.* 2019;25(1):24–9.
5. Foulquier N, Redou P, Le Gal C, Rouvière B, Pers J-O, Saraux A. Pathogenesis-based treatments in primary Sjogren's syndrome using artificial intelligence and advanced machine learning techniques: a systematic literature review. *Hum Vaccines Immunother.* 2018;17:1–6.
6. Ghahramani Z. Probabilistic machine learning and artificial intelligence. *Nature.* 2015;521(7553):452–9.
7. Luo D, Wang P, Lu F, Elias J, Sparks JA, Lee YC. Mobile apps for individuals with rheumatoid arthritis: a systematic review. *JCR J Clin Rheumatol.* 2018;25:133–41.
8. Saraux L, Devauchelle-Pensec V, Saraux A. Plea for standardisation of disease activity scores. *Rheumatology.* 2019. <https://doi.org/10.1093/rheumatology/kez117>.
9. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. 2012. <http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-network>. Accessed 19 Mar 2019.
10. Üreten K, Erbay H, Maraş HH. Detection of rheumatoid arthritis from hand radiographs using a convolutional neural network. *Clin Rheumatol.* 2019. <http://link.springer.com/10.1007/s10067-019-04487-4>. Accessed 19 Mar 2019.