Title	Automatic evaluation of atlantoaxial subluxation in rheumatoid arthritis by a deep learning model
Author(s)	Okita, Yasutaka; Hirano, Toru; Wang, Bowen et al.
Citation	Arthritis Research and Therapy. 2023, 25(1), p. 181
Version Type	VoR
URL	https://hdl.handle.net/11094/93229
rights	This article is licensed under a Creative Commons Attribution 4.0 International License.
Note	

Osaka University Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

Osaka University

RESEARCH ARTICLE

Automatic evaluation of atlantoaxial subluxation in rheumatoid arthritis by a deep learning model

Yasutaka Okita[1]* , Toru Hirano[2], Bowen Wang[3], Yuta Nakashima[3], Saki Minoda[1], Hajime Nagahara[3] and Atsushi Kumanogoh[1,4,5]

Abstract

Background This work aims to develop a deep learning model, assessing atlantoaxial subluxation (AAS) in rheumatoid arthritis (RA), which can often be ambiguous in clinical practice.

Methods We collected 4691 X-ray images of the cervical spine of the 906 patients with RA. Among these images, 3480 were used for training the deep learning model, 803 were used for validating the model during the training process, and the remaining 408 were used for testing the performance of the trained model. The two-dimensional key points' detection model of Deep High-Resolution Representation Learning for Human Pose Estimation was adopted as the base convolutional neural network model. The model inferred four coordinates to calculate the atlantodental interval (ADI) and space available for the spinal cord (SAC). Finally, these values were compared with those by clinicians to evaluate the performance of the model.

Results Among the 408 cervical images for testing the performance, the trained model correctly identified the four coordinates in 99.5% of the dataset. The values of ADI and SAC were positively correlated among the model and two clinicians. The sensitivity of AAS diagnosis with ADI or SAC by the model was 0.86 and 0.97 respectively. The specificity of that was 0.57 and 0.5 respectively.

Conclusions We present the development of a deep learning model for the evaluation of cervical lesions of patients with RA. The model was demonstrably shown to be useful for quantitative evaluation.

Keywords Rheumatoid arthritis, Deep learning, Machine learning, Cervical spine, Atlantoaxial subluxation

*Correspondence:
Yasutaka Okita
y.okita@imed3.med.osaka-u.ac.jp
[1] Department of Respiratory Medicine and Clinical Immunology, Osaka University Graduate School of Medicine, 2-2, Yamadaoka, Suita, Osaka 565-0871, Japan
[2] Department of Rheumatology, Nishinomiya Municipal Central Hospital, Hyogo, Japan
[3] Osaka University Institute for Datability Science (IDS), Suita, Osaka, Japan
[4] Laboratory of Immunopathology, World Premier International Immunology Frontier Research Center, Osaka University, Suita, Osaka, Japan
[5] The Institute for Open and Transdisciplinary Research Initiatives (OTRI), Osaka, Japan

Background

Rheumatoid arthritis (RA) is a chronic autoimmune disease characterized by synovitis and multiple joint destruction. The spine is affected in addition to the joints of extremities in patients with RA. Cervical spine lesion is the most frequent spinal lesion due to RA and it is found in 25–80% of patients with RA [1–10]. The frequency of cervical spine lesions in patients with RA is not necessarily decreasing [11, 12], even though many patients achieve remission or low disease activity [13]. One of the reasons is that the patients with RA have been aging [14] and the cervical

Okita *et al. Arthritis Research & Therapy*   (2023) 25:181

Page 2 of 7

spine lesions are also affected by aging. Cervical spine lesions are classified into two types, upper and middle/lower cervical spine lesions, with upper cervical spine lesions being more common [12]. Cervical spine lesion causes not only neck pain but also vascular compression [15], radiculopathy, and myelopathy because of cervical spine instability, which can be life-threatening [16]. Thus, patients with cervical spine lesions require careful attention in the management of RA. Screening and reproducible evaluation by X-ray examination of atlantoaxial subluxation (AAS) of the upper cervical spine are important, because it is the most common type [17] and a risk for complications of other types of cervical lesions such as vertical subluxation of the axis and subaxial subluxation [5]. AAS is diagnosed with an atlantodental interval (ADI) of more than 3 mm or the space available for the spinal cord (SAC) of 13 mm or less [18]. However, the cervical subluxation is often overlooked for several reasons. Firstly, half of patients with radiographic instability are asymptomatic [3]. Secondly, typical neurological abnormalities such as hyperreflexia are not observed due to peripheral joint destruction. Thirdly, radiographic evidence of AAS sometimes appears to be decreased in X-ray images due to the development of basilar invagination (pseudostabilization) [11].

Recently, artificial intelligence (AI) has been developed and applied in medicine such as the diagnosis of diabetic retinopathy [19] and support for endoscopic diagnosis [20]. Convolutional neural network (CNN), which is an algorithm suitable for image processing among AI, shows remarkable performance in classifying various images, including the diagnosis of skin cancer [21, 22], the detection of lymph node metastases in women with breast cancer [23], classification of interstitial lung diseases [24], and analysis of genomic data [25]. As an additional merit, AI can process large amounts of data in a short time with good reproducibility. In the field of RA, CNNs have been developed for several uses, such as the diagnosis of RA [26], and evaluation of synovitis by ultrasound images [27, 28]. Previously, we developed CNN for the analysis of finger joints by X-ray imaging, which is the gold standard for the evaluation of joint destruction due to RA [29].

In this work, we developed a deep learning model to evaluate AAS on cervical spine X-ray images of patients with RA and evaluated its performance. The identification of four coordinates is mandatory for ADI and SAC calculations. The two-dimensional key points' detection model of Deep High-Resolution Representation Learning for Human Pose Estimation (HRNet) [30] was adopted as the base CNN model.

Methods

Patients and images

Nine hundred and thirty-six patients diagnosed with RA based on the American Rheumatism Association 1987 revised criteria [31] or the 2010 ACR-EULAR RA classification criteria [32] were enrolled. All patients received treatment at Osaka University Hospital and were enrolled in the institute's cohort of patients with RA. Clinical information, such as age, gender, and disease duration, was collected from medical records. The present work was approved by the Ethics Committee of Osaka University Hospital and was conducted according to the Declaration of Helsinki. The board waived the requirement for patient informed consent by posting the opt-out information on the hospital's home page.

We retrospectively collected digital X-ray images of neutral position of the right lateral cervical spine and anteflexion/retroflexion positions if they had been taken, that were taken from January 1, 2016, to May 31, 2021. The images of 21 patients were excluded because of a history of cervical spine surgery as well as those of nine additional patients because of extremely severe destruction of the cervical spine, in which the atlantoaxial joint could not be identified. Finally, 4691 images of the 906 patients were used for machine learning.

The nine hundred and six patients were given serial numbers and the numbers were randomly sorted, by numpy.ramdom.permutation function in NumPy 1.21.6. The patients from the beginning to the 680th in the sequence were selected for the training dataset, the patients from the 681st to 820th were selected for the training dataset, and the remaining patients were selected for the test dataset. Finally, a training dataset to adjust network parameters included 3480 images of 680 patients, a validation dataset to monitor the training process included 803 images of 140 patients, and a test dataset to assess the trained model included 408 images of 86 patients.

Machine learning step

Based on the consensus of two rheumatologists (Y.O. and S.M.), all images were labeled with four coordinates essential for ADI and SAC calculation and Digital Imaging and Communications in Medicine (DICOM) tag (0028,0030) meaning pixel spacing (length per pixel). The machine learning program was written using Python 3.6.9, NumPy 1.24.3, OpenCV-Python 4.1.2.30, and

Okita *et al. Arthritis Research & Therapy*   (2023) 25:181

Page 3 of 7

Table 1 Characteristics of the patients

Characteristic	Total
n	906
Sex, female/male	706 / 200
Age, years, median (interquartile ranges)	66 (54, 73)
Disease duration, years, median (interquartile ranges)	8 (1, 17)
RF positive, %	79.0%
ACPA positive, %	63.0%
Number of radiographs	4691

RF rheumatoid factor, *ACPA* anti-citrullinated peptide antibody

Matplotlib 3.2.2. The two-dimensional key points' detection model of HRNet which is able to maintain high-resolution representations through the whole process and give great performance was adopted as the base CNN model. It was combined with various modules (transfer learning and ADI and SAC calculations) to improve performance. The batch size for training was 32.

Model performance evaluation

Using the test dataset, we examined the performance of the trained model by comparing the ADI and SAC assessed by the model to those by the two clinicians. Sensitivity and specificity were calculated when the judgment of clinician 1 was defined as true; that is, an $ADI \geq 3$ mm or $SAC \leq 13$ mm was determined as AAS positive. Differences in ADI and SAC among the model, clinician 1,
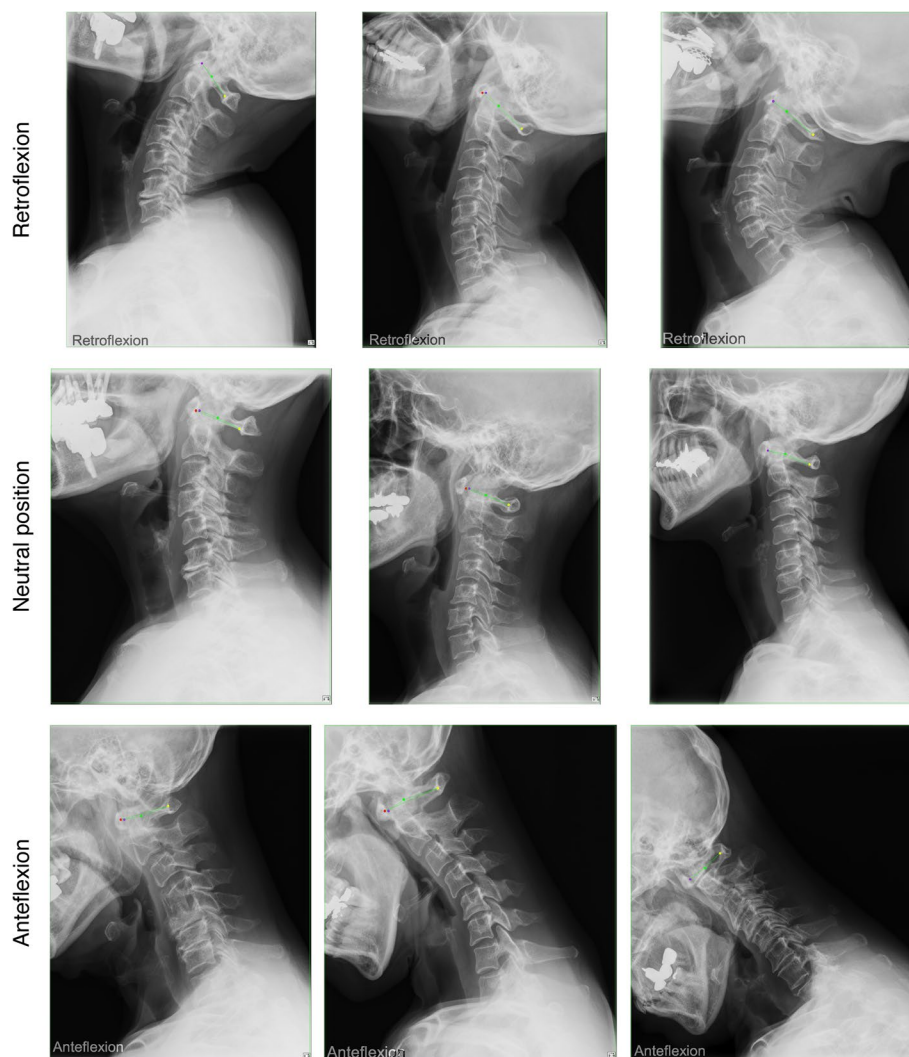

Fig. 1 Detection of four coordinates for atlantodental interval and space available for the spinal cord. Representative images processed by the model. The four coordinates mandatory for atlantodental interval and space available for the spinal cord calculations detected by the model are shown in red, purple, green, and yellow dots from the ventral side

Okita *et al. Arthritis Research & Therapy*   (2023) 25:181

Page 4 of 7

and clinician 2 were compared and statistically examined using an equivalence test with a hypothesis value of 0 and a range of 1.5 mm for ADI or 2 mm for SAC. Statistical analyses were conducted using JMP software (Pro 16.0.0).

Results

Patients and images

Table 1 shows the characteristics of the 906 patients. The proportion of females was 77.9%. The median and interquartile range of age and disease duration at the time the first X-ray image was obtained were 66 years (54–73 years) and 8 years (1–17 years), respectively. The seropositivity of rheumatoid factor was 79.0% and that of anti-citrullinated peptide antibody was 63.0%. The profiles of patients indicated that the participants enrolled in this work were representative of the general RA population. The values of ADI and SAC measured by clinicians are shown in sFig. 1. The median and interquartile range of ADI and SAC were 1.61 mm (1.00–2.92 mm) and 19.81 mm (17.23–21.52 mm), respectively.
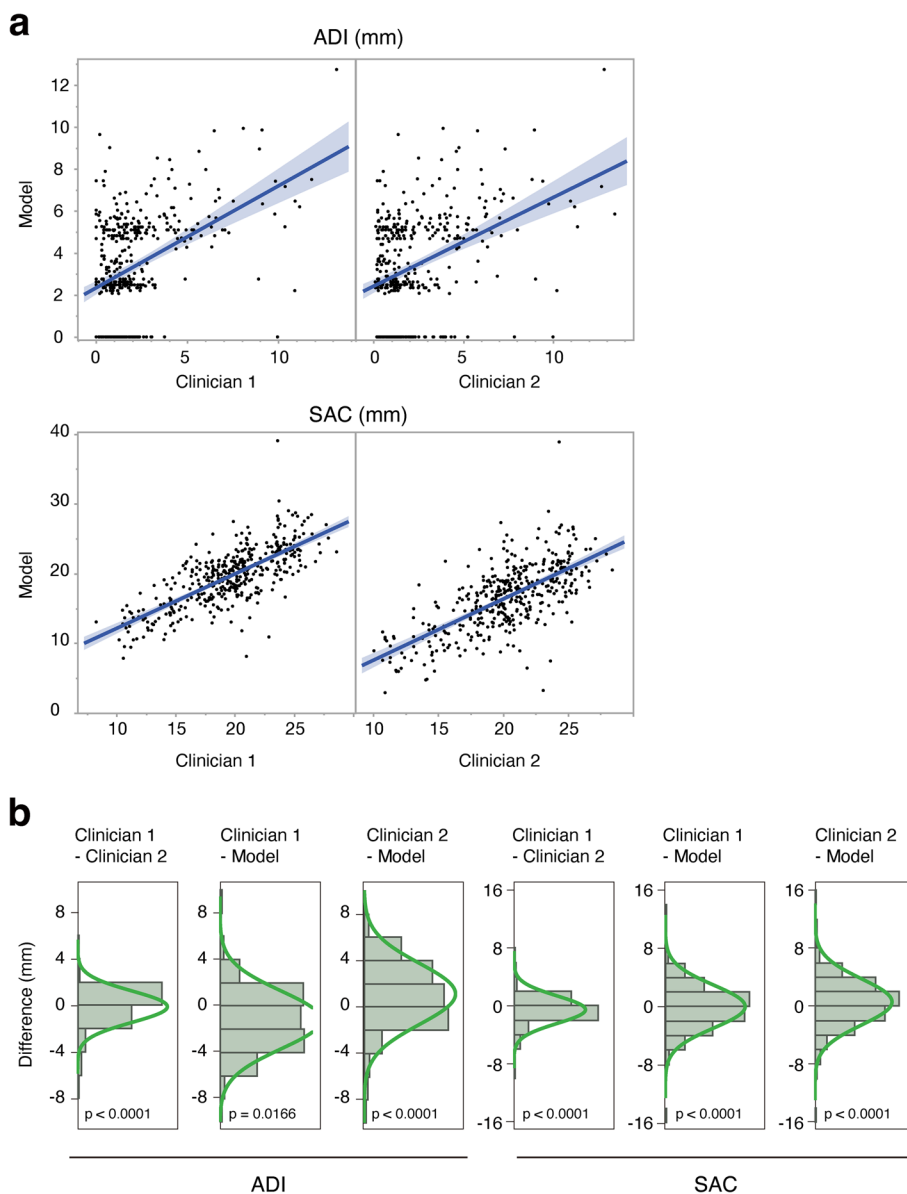

Fig. 2 Evaluation of model performance. **a** Correlations of values of atlantodental interval (ADI, upper column) or space available for the spinal cord (SAC, lower column) among the model, clinician 1, and clinician 2. **b** Distribution of the difference of ADI (left three columns) and SAC (right three columns) among the model, clinician 1, and clinician 2. Each graph is scaled by the range of ADI or SAC. The values were statistically examined by equivalence test with hypothesis value zero with a range of ± 1.5 mm for ADI and ± 2.0 mm for SAC

Okita *et al. Arthritis Research & Therapy* (2023) 25:181

Page 5 of 7

Key points' detection

As shown in Fig. 1a, the four coordinates detected by the model were shown in red, purple, green, and yellow dots from the ventral side. The heatmap loss value and Percentage of Correct Keypoints (PCK) were less than 0.0001 and 0.2551 at 40 epochs, respectively. The mean distance between true and estimated coordinates is 2.66 mm (standard deviation 4.98 mm). The distance was more than 10 mm in two images, suggesting an estimated probability of 99.5%. These two images were judged to be outliers.

Evaluation of model performance

As shown in Fig. 2a, the values of ADI were positively correlated among the model, clinician 1, and clinician 2 (the correlation coefficients were 0.43 and 0.40 respectively). The values of SAC were also positively correlated among the model, clinician 1, and clinician 2 (the correlation coefficients were 0.73 and 0.68 respectively). The differences in ADI or SAC among the model, clinician 1, and clinician 2 are shown in Fig. 2b. These differences were normally distributed and could be interpreted as zero with a range of ± 1.5 mm for ADI and ± 2.0 mm for SAC by equivalence test.

The sensitivity of AAS diagnosis with ADI by the model was 0.86 and the specificity was 0.57 (Table 2). The sensitivity of AAS diagnosis with SAC by the model was 0.97 and the specificity was 0.5 (Table 2).

Discussion

We developed a deep learning model for atlantoaxial joint detection and ADI/SAC evaluation, using images of the cervical spine from patients with RA. While developing the machine learning model, we created and evaluated several algorithms such as programs to formulate the key points' detection as a CNN-based regression problems. Consequently, this model which is able

to maintain high-resolution representations through the whole process and gave the best performance was reported.

In daily clinical practice, it is often difficult to diagnose cervical spine lesions for several reasons. Half of patients with cervical spine lesions are asymptomatic [3]. Deep tendon reflexes are not often exaggerated because of limb joint destruction in patients with RA. There are some cases where radiographic evidence of AAS sometimes appear to be decreased in X-ray image due to the development of basilar invagination. This pattern has been described as "pseudostabilization." Therefore, cervical spine lesions can be discovered at advanced stage such as the Ranawat classification Class IIIB [33, 34]. Our model could diagnose AAS with a sensitivity of 0.97 and specificity of 0.57. Our findings show that the developed AI-based model can accurately and rapidly identify AAS, thereby potentially reducing oversight of AAS in patients in clinical settings.

Cervical spine lesion in RA is progressive, and spinal cord symptoms may result in quadriplegia or shortened lifespan [35]. Thus, it is necessary to evaluate disease progression over time using a reliable approach to ensure timely treatment for the suppression of cervical spine destruction. In addition to identifying AAS, our model can be also useful in assessing the interval change of cervical spine lesions in patients with RA over time as the outputs of ADI and SAC are numerical values.

Our work has a few limitations. First, X-ray images used in this study were retrospectively collected. Prospective studies are needed to verify whether our model is really useful in assessing changes in cervical spine lesions in RA patients over time. Second, X-ray images used in this study were taken at Osaka University Hospital. Further machine learning with images taken at other hospitals is required to improve the validity and robustness of the model. Third, our model can assess AAS, but it cannot assess vertical subluxation of the axis or subaxial subluxation, both of which are also found in patients with RA. As a next step, it is considered necessary to develop AI to evaluate them. Fourth, our model could not detect the four coordinates in two images. In one of them, the patient bent forward until almost horizontal.

Conclusions

We introduced a deep learning model using CNN to assess atlantoaxial joint destruction in patients with RA. This model provides a partial assessment of the many joints that can be destroyed in RA. The introduction of AI in healthcare can be useful for oversight prevention, time-savings, reductions in effort, health surveys, and assessments by healthcare professionals, both specialists

Table 2 Diagnosis of atlantoaxial subluxation

ADI		By the model	
		≥ 3 mm	< 3 mm
By clinician 1	≥ 3 mm	68	11
	< 3 mm	141	186
By clinician 2	≥ 3 mm	83	23
	< 3 mm	126	174
SAC		By the model	
		≥ 13 mm	< 13 mm
By clinician 1	≥ 13 mm	367	11
	< 13 mm	14	14
By clinician 2	≥ 13 mm	372	14
	< 13 mm	9	11

ADI atlantodental interval, *SAC* space available for the spinal cord

Okita *et al. Arthritis Research & Therapy*   (2023) 25:181

Page 6 of 7

and non-specialists. Importantly, our methodology can be applied to other joints, such as the elbow, shoulder, hip, knee, foot, and spine, as well as for other conditions or disorders, including osteoporosis, fracture, and bone tumor. We conclude that image processing with a trained CNN model is a promising tool to assess radiographic cervical destruction in RA.

Abbreviations

AAS	Atlantoaxial subluxation
ADI	Atlantodental interval
AI	Artificial intelligence
CNN	Convolutional neural network
RA	Rheumatoid arthritis
SAC	Space available for the spinal cord

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13075-023-03172-x.

> **Additional file 1: Sup Fig. 1.** Distribution of ADI and SAC. **a** Atlantodental interval (ADI, mm). **b** Space available for the spinal cord (SAC, mm).

Authors' contributions
Y.O. and T.H. designed the study and performed the statistical analyses. Y.O., B.W., and Y.N. described the program. Y.O. and S.M. measured ADI and SAC. All authors wrote and revised the manuscript.

Availability of data and materials
The datasets used and analyzed during the current work are available from the corresponding author upon reasonable request. The code is available at https://github.com/ya-o/neck_RA.

Declarations

Ethics approval and consent to participate
This work was approved by the Ethics Committee of Osaka University Hospital and was conducted according to the Declaration of Helsinki. The board waived the requirement for patient informed consent by posting the opt-out information in the hospital's home page.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

References

1. Smith PH, Benn RT, Sharp J. Natural history of rheumatoid cervical luxations. Ann Rheum Dis. 1972;31(6):431–9.
2. Cabot A, Becker A. The cervical spine in rheumatoid arthritis. Clin Orthop Relat Res. 1978;131:130–40.
3. Pellicci PM, Ranawat CS, Tsairis P, Bryan WJ. A prospective study of the progression of rheumatoid arthritis of the cervical spine. J Bone Joint Surg Am. 1981;63(3):342–50.
4. Morizono Y, Sakou T, Kawaida H. Upper cervical involvement in rheumatoid arthritis. Spine (Phila Pa 1976). 1987;12(8):721–5.
5. Rana NA. Natural history of atlanto-axial subluxation in rheumatoid arthritis. Spine (Phila Pa 1976). 1989;14(10):1054–6.
6. Rajangam K, Thomas IM. Frequency of cervical spine involvement in rheumatoid arthritis. J Indian Med Assoc. 1995;93(4):138–9, 137.
7. Oda T, Fujiwara K, Yonenobu K, Azuma B, Ochi T. Natural course of cervical spine lesions in rheumatoid arthritis. Spine (Phila Pa 1976). 1995;20(10):1128–35.
8. Paimela L, Laasonen L, Kankaanpää E, Leirisalo-Repo M. Progression of cervical spine changes in patients with early rheumatoid arthritis. J Rheumatol. 1997;24(7):1280–4.
9. Fujiwara K, Yonenobu K, Ochi T. Natural history of upper cervical lesions in rheumatoid arthritis. J Spinal Disord. 1997;10(4):275–81.
10. Taniguchi D, Tokunaga D, Hase H, Mikami Y, Hojo T, Ikeda T, Oda R, Takatori R, Imai K, Kida Y, et al. Evaluation of lateral instability of the atlanto-axial joint in rheumatoid arthritis using dynamic open-mouth view radiographs. Clin Rheumatol. 2008;27(7):851–7.
11. Cha TD, An HS. Cervical spine manifestations in patients with inflammatory arthritides. Nat Rev Rheumatol. 2013;9(7):423–32.
12. Yurube T, Sumi M, Nishida K, Miyamoto H, Kohyama K, Matsubara T, Miura Y, Sugiyama D, Doita M, Conference KS. Incidence and aggravation of cervical spine instabilities in rheumatoid arthritis: a prospective minimum 5-year follow-up study of patients initially without cervical involvement. Spine (Phila Pa 1976). 2012;37(26):2136–44.
13. Aga AB, Lie E, Uhlig T, Olsen IC, Wierød A, Kalstad S, Rødevand E, Mikkelsen K, Kvien TK, Haavardsholm EA. Time trends in disease activity, response and remission rates in rheumatoid arthritis during the past decade: results from the NOR-DMARD study 2000–2010. Ann Rheum Dis. 2015;74(2):381–8.
14. Kato E, Sawada T, Tahara K, Hayashi H, Tago M, Mori H, Nishino J, Matsui T, Tohma S. The age at onset of rheumatoid arthritis is increasing in Japan: a nationwide database study. Int J Rheum Dis. 2017;20(7):839–45.
15. Delamarter RB, Bohlman HH. Postmortem osseous and neuropathologic analysis of the rheumatoid cervical spine. Spine (Phila Pa 1976). 1994;19(20):2267–74.
16. Riise T, Jacobsen BK, Gran JT. High mortality in patients with rheumatoid arthritis and atlantoaxial subluxation. J Rheumatol. 2001;28(11):2425–9.
17. Dreyer SJ, Boden SD. Natural history of rheumatoid arthritis of the cervical spine. Clin Orthop Relat Res. 1999;366:98–106.
18. Boden SD, Dodge LD, Bohlman HH, Rechtine GR. Rheumatoid arthritis of the cervical spine. A long-term analysis with predictors of paralysis and recovery. J Bone Joint Surg Am. 1993;75(9):1282–97.
19. van der Heijden AA, Abramoff MD, Verbraak F, van Hecke MV, Liem A, Nijpels G. Validation of automated screening for referable diabetic retinopathy with the IDx-DR device in the Hoorn Diabetes Care System. Acta Ophthalmol. 2018;96(1):63–8.
20. Yamada M, Saito Y, Imaoka H, Saiko M, Yamada S, Kondo H, Takamaru H, Sakamoto T, Sese J, Kuchiba A, et al. Development of a real-time endoscopic image diagnosis support system using deep learning technology in colonoscopy. Sci Rep. 2019;9(1):14465.
21. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017;542(7639):115–8.
22. Han SS, Moon IJ, Lim W, Suh IS, Lee SY, Na JI, Kim SH, Chang SE. Keratinocytic skin cancer detection on the face using region-based convolutional neural network. JAMA Dermatol. 2020;156(1):29–37.
23. Ehteshami Bejnordi B, Veta M, Johannes van Diest P, van Ginneken B, Karssemeijer N, Litjens G, van der Laak JAWM, Hermsen M, Manson QF, Balkenhol M et al: Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. JAMA. 2017;318(22):2199–2210.

Okita *et al. Arthritis Research & Therapy*   (2023) 25:181

Page 7 of 7

24. Anthimopoulos M, Christodoulidis S, Ebner L, Christe A, Mougiakakou S. Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. IEEE Trans Med Imaging. 2016;35(5):1207–16.

25. Poplin R, Chang PC, Alexander D, Schwartz S, Colthurst T, Ku A, Newburger D, Dijamco J, Nguyen N, Afshar PT, et al. A universal SNP and small-indel variant caller using deep neural networks. Nat Biotechnol. 2018;36(10):983–7.

26. Fukae J, Isobe M, Hattori T, Fujieda Y, Kono M, Abe N, Kitano A, Narita A, Henmi M, Sakamoto F, et al. Convolutional neural network for classification of two-dimensional array images generated from clinical information may support diagnosis of rheumatoid arthritis. Sci Rep. 2020;10(1):5648.

27. Christensen ABH, Just SA, Andersen JKH, Savarimuthu TR. Applying cascaded convolutional neural network design further enhances automatic scoring of arthritis disease activity on ultrasound images from rheumatoid arthritis patients. Ann Rheum Dis. 2020;79(9):1189–93.

28. Andersen JKH, Pedersen JS, Laursen MS, Holtz K, Grauslund J, Savarimuthu TR, Just SA. Neural networks for automatic scoring of arthritis disease activity on ultrasound images. RMD Open. 2019;5(1):e000891.

29. Hirano T, Nishide M, Nonaka N, Seita J, Ebina K, Sakurada K, Kumanogoh A. Development and validation of a deep-learning model for scoring of radiographic finger joint destruction in rheumatoid arthritis. Rheumatol Adv Pract. 2019;3(2):rkz047.

30. Sun K, Xiao B, Liu D, Wang J. Deep High-Resolution Representation Learning for Human Pose Estimation. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2019. p. 5693–703.

31. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, Healey LA, Kaplan SR, Liang MH, Luthra HS. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum. 1988;31(3):315–24.

32. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO, Birnbaum NS, Burmester GR, Bykerk VP, Cohen MD, et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Ann Rheum Dis. 2010;69(9):1580–8.

33. Ranawat CS, O'Leary P, Pellicci P, Tsairis P, Marchisello P, Dorr L. Cervical spine fusion in rheumatoid arthritis. J Bone Joint Surg Am. 1979;61(7):1003–10.

34. Casey AT, Crockard HA, Bland JM, Stevens J, Moskovich R, Ransford AO. Surgery on the rheumatoid cervical spine for the non-ambulant myelopathic patient-too much, too late? Lancet. 1996;347(9007):1004–7.

35. Sunahara N, Matsunaga S, Mori T, Ijiri K, Sakou T. Clinical course of conservatively managed rheumatoid arthritis patients with myelopathy. Spine (Phila Pa 1976). 1997;22(22):2603–7; discussion 2608.

Publisher's Note