

The differential diagnosis of IgG4-related disease based on machine learning

Motohisa Yamamoto (✉ mocha@ims.u-tokyo.ac.jp)

IMSUT Hospital: Tokyo Daigaku Ikagaku Kenkyujo Fuzoku Byoin

Masanori Nojima

IMSUT Hospital: Tokyo Daigaku Ikagaku Kenkyujo Fuzoku Byoin

Ryuta Kamekura

Sapporo Medical University: Sapporo Ika Daigaku

Akiko Kuribara-Souta

IMSUT Hospital: Tokyo Daigaku Ikagaku Kenkyujo Fuzoku Byoin

Masaaki Uehara

IMSUT Hospital: Tokyo Daigaku Ikagaku Kenkyujo Fuzoku Byoin

Hiroki Yamazaki

IMSUT Hospital: Tokyo Daigaku Ikagaku Kenkyujo Fuzoku Byoin

Noritada Yoshikawa

IMSUT Hospital: Tokyo Daigaku Ikagaku Kenkyujo Fuzoku Byoin

Satsuki Aochi

IMSUT Hospital: Tokyo Daigaku Ikagaku Kenkyujo Fuzoku Byoin

Ichiro Mizushima

Kanazawa University: Kanazawa Daigaku

Takayuki Watanabe

Shinshu University Graduate School of Medicine School of Medicine: Shinshu Daigaku Daigakuin
Igaku-kei Kenkyu-ka Igakubu

Aya Nishiwaki

University of Tokyo: Tokyo Daigaku

Toshihiko Komai

University of Tokyo: Tokyo Daigaku

Hirofumi Shoda

University of Tokyo: Tokyo Daigaku

Koji Kitagori

Kyoto University: Kyoto Daigaku

Hajime Yoshifuji

Kyoto University: Kyoto Daigaku

Hideaki Hamano

Nagano Prefectural Kiso Hospital

Mitsuhiro Kawano

Kanazawa University: Kanazawa Daigaku

Ken-ichi Takano

Sapporo Medical University: Sapporo Ika Daigaku

Keishi Fujio

University of Tokyo: Tokyo Daigaku

Hirotohi Tanaka

IMSUT Hospital: Tokyo Daigaku Ikagaku Kenkyujo Fuzoku Byoin

Research article

Keywords: artificial intelligence, differential diagnosis, IgG4-related disease, machine learning

Posted Date: October 26th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-968651/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Arthritis Research & Therapy on March 19th, 2022. See the published version at <https://doi.org/10.1186/s13075-022-02752-7>.

Abstract

Introduction:

To eliminate the disparity and maldistribution of physicians and medical specialty services, the development of diagnostic support for rare diseases using artificial intelligence is being promoted. Immunoglobulin G4 (IgG4)-related disease (IgG4-RD) is a rare disorder often requiring special knowledge and experience to diagnose. In this study, we investigated the possibility of differential diagnosis of IgG4-RD based on basic patient characteristics and blood test findings using machine learning.

Methods

Six-hundred and two patients with IgG4-RD and 212 patients with non-IgG4-RD that needed to be differentiated who visited the participating institutions were included in the study. Ten percent of the subjects were randomly excluded as a validation sample. Among the remaining cases, 80% were used as training samples, and the remaining 20% were used as test samples. Finally, validation was performed on the validation sample. The analysis was performed using a decision tree and a random forest model. Furthermore, a comparison was made between conditions with and without the serum IgG4 concentration. Accuracy was evaluated using the area under the receiver-operating characteristic (AUROC) curve.

Results

In diagnosing IgG4-RD, AUROC curve values of the decision tree and the random forest method were 0.905 and 0.970, respectively, when serum IgG4 levels were included in the analysis. Excluding serum IgG4 levels, the AUROC curve value of the analysis by the random forest method was 0.919.

Conclusion

Based on machine learning in a multicenter collaboration, with or without serum IgG4 data, basic patient characteristics and blood test findings alone were sufficient to differentiate IgG4-RD from non-IgG4-RD.

Introduction

In the current severe acute respiratory syndrome coronavirus 2 pandemic, disparities in medical care have become a problem. The application of artificial intelligence (AI) in medicine has been promoted for a long time to solve regional disparities and uneven distribution of physicians and medical specialty services. However, we now have to accelerate the pace of its application. We are developing the use of AI for diagnosing rare diseases wherever the patients live. The basic technology behind AI is a machine learning algorithm, applied in various ways in the biomedical field [1]. However, not much progress has been made

in the field of rheumatology. In diagnostic imaging, many studies have been conducted to detect early-stage lesions using AI with joint sonography and magnetic resonance imaging (MRI), leading to the diagnosis of rheumatoid arthritis (RA) [2, 3].

Rheumatic diseases are currently diagnosed using patterned diagnostic criteria based on a combination of physical, hematological, and imaging findings. In addition, the therapeutic strategy for rheumatic diseases is decided after carefully considering the distribution and degree of disability. These are the areas in which AI excels the most. Currently, the diagnosis of IgG4-related disease (IgG4-RD) is based on blood test results; findings from imaging examinations such as computed tomography (CT), MRI, and fluorodeoxyglucose positron emission tomography (FDG-PET), and histopathological findings. As a result, the invasiveness to the patients is high, and the high cost of medical care has become a problem. Thus, this study investigated whether AI can be trained to differentiate IgG4-RD from other rheumatic diseases by learning the typical cases of both IgG4-RD and non-IgG4-RD, and whether proper diagnosis is possible. The results of this study are expected to be useful for assisting nonspecialist physicians in the community make appropriate diagnosis and treatment decisions for patients with IgG4-RD.

Methods

Patients

The subjects were 602 patients with IgG4-RD who visited the Institute of Medical Science, The University of Tokyo (IMSUT) Hospital, Kanazawa University Hospital, Shinshu University Hospital, Kyoto University Hospital, and Sapporo Medical University Hospital between April 1997 and June 2019 and provided consent to participate in the study based on the information provided. In addition, 212 patients with other rheumatic disorders that need to be differentiated from IgG4-RD who visited the aforementioned institutions between January 2019 and June 2019 were included in the study. IgG4-RD was diagnosed based on the comprehensive diagnostic criteria for IgG4-RD (Ministry of Health, Labour and Welfare Research Group, Japan, 2011) [4]. Among those who visited the aforementioned institutions before 2010, only those with a definitive diagnosis according to the comprehensive diagnostic criteria were included in this study. Concerning the diagnosis of other rheumatic diseases, the diagnostic criteria developed by the relevant societies and research groups were used [5–9]. The diagnostic criteria that were used for microscopic polyangiitis are shown in Supplementary Table 1. Both multicentric Castleman's disease and sarcoidosis were diagnosed clinically, including histopathological findings.

Machine Learning And Statistical Analyses

First, a dataset including two basic patient characteristics and 29 laboratory findings was created for each case at the first visit (Table 1). First, 10% of the patients and controls were randomly selected as a validation sample. Of the remaining cases, 80% and 20% were used as training and test samples, respectively. This process was performed randomly by the computer in each of IgG4-RD and non-IgG4-RD

groups. Finally, validation was performed on the validation sample, initially separated. A prediction model was developed using the training sample so that the outcome was set to extract IgG4-RD cases from among the cases of diseases that need to be differentiated from IgG4-RD. We also performed analyses for situations with known and unknown serum IgG4 concentration, which are considered to be important variables in the diagnosis of IgG4-RD. In the latter situation, we purposely excluded the IgG4 concentration from the analysis. The accuracy of the model was retrospectively evaluated by drawing a receiver-operating characteristic (ROC) curve from the sensitivity and specificity of the validation sample and calculating the area under the curve. Missing data were imputed for by the median of each disease group. R version 3.6.1 software (downloaded at <https://cran.ism.ac.jp>) was used. The following regression tree (CART) and random forest models were applied using the R package, “rpart”(v. 4.1-15) (https://cran.ism.ac.jp/bin/macosx/contrib/4.0/rpart_4.1-15.tgz) and “randomForest”(v. 4.6-14) (https://cran.ism.ac.jp/bin/macosx/contrib/4.0/randomForest_4.6-14.tgz), respectively. In terms of model fitting, when the serum IgG4 values were known, the CART method used a cp (a parameter indicating the complexity of the tree model) value of 0.029, the Random Forest method used a mtry (the number of variables to be employed in the model) of 10, and the ntree (the number of decision trees to be tried) of 300, and when the serum IgG4 values were unknown, the For the CART method, the cp value was 0,017, and for the Random Forest method, the mtry was 5 and the ntree was 600 for optimization. Intergroup comparisons were performed using a two-tailed *t*-test. *P* values less than 0.05 were used to denote significance.

Results

Patients' profiles

Among the 602 patients with IgG4-RD, 350 were male and 252 were female (male-female ratio, 1.4:1). The mean (\pm standard deviation (SD)) age at the first visit was 64.11 ± 11.46 years. The mean (\pm SD) optical density serum IgG concentration was $2,204.14 \pm 1,146.86$ mg/dL, and the mean (\pm SD) serum IgG4 concentration was 666.00 ± 660.46 mg/dL. Of the 602 patients, 499 had dacryoadenitis and sialadenitis, 192 had autoimmune pancreatitis, 51 had sclerosing cholangitis, 107 had renal involvement, 87 had respiratory involvement, and 182 had retroperitoneal fibrosis (including duplicates). The number of patients with two or more organ lesions was 350 (58.1%) (Table 2).

In the 212 patients with the diseases that needed to be differentiated from IgG4-RD, the male-female ratio was 1:4.0 (42 men and 170 women). The mean (\pm SD) age at the first visit was 54.20 ± 16.09 years, which was significantly younger than that in the IgG4-RD group ($p < 0.001$). The details of the diseases are shown in Table 2. Of the patients with non-IgG4-RD, 114 had Sjögren's syndrome (SS) alone, and 25 had systemic lupus erythematosus (SLE). In addition, 19 patients with multicentric Castleman's disease (MCD) and 11 patients with eosinophilic granulomatosis with polyangiitis (EGPA), which often presents with elevated serum IgG4 levels, were included. In the MCD group, the mean (\pm SD) serum IgG and IgG4 concentrations were $3,828.05 \pm 1,850.66$ and $710.42 \pm 1,156.82$ mg/dL, respectively. No significant

difference in serum IgG4 level was found between the MCD and IgG4-RD groups ($p = 0.63$). The missing data rate was less than 3%.

Prediction of IgG4-RD diagnosis in patients with rheumatic diseases requiring differentiation

We performed diagnostic prediction of IgG4-RD using 814 cases of IgG4-RD and non-IgG4-RD. When the serum IgG4 level was known, the diagnosis of IgG4-RD was predicted by a decision tree (Figure 1A). First, if the serum IgG4 level is ≥ 115 mg/dL, go to the right, and if the CRP is < 2.08 mg/dL, go to the rightmost column, which is probable IgG4-RD. If CRP ≥ 2.08 mg/dL, but CH50 < 51.5 U/mL, the diagnosis of IgG4-RD is more likely. If the laboratory findings meet all of the following conditions, IgG4 < 115 mg/dL, C3 ≥ 96.5 mg/dL, sIL-2R < 671.5 U/mL, and IgG < 1843.5 mg/dL, the likelihood of IgG4-RD increases significantly. The ROC curve for this algorithm is shown in Figure 1B. The accuracy, sensitivity, and specificity of the algorithm were 0.932, 0.981, and 0.810, respectively, and the AUC was 0.937. The validation of this algorithm showed that its accuracy, sensitivity, and specificity were 0.915, 0.983, and 0.727, respectively, and the AUC was 0.905 (Figure 1C).

The same subjects were then analyzed using the random forest method. The Gini Impurity is shown in Figure 2A. In this algorithm, the serum IgG4 concentration is the most important variable, followed by the serum sIL-2R concentration, age at the first visit, serum IgM concentration, and serum CRP concentration. The ROC curve for this algorithm is shown in Figure 2B. The accuracy, sensitivity, and specificity of the algorithm were 0.934, 0.981, and 0.833, respectively, and the AUC was 0.993. The validation of this algorithm showed that its accuracy, sensitivity, and specificity were 0.951, 0.983, and 0.864, respectively, and the AUC was 0.970 (Figure 2C).

When the serum IgG4 level was unknown, the diagnosis of IgG4-RD was predicted by a decision tree (Figure 3A). First, if the serum IgM level is < 139.5 mg/dL, go to the right, and if the IgG level is $\geq 1,254$ mg/dL and age at the first visit ≥ 49.5 years, go to the rightmost column, which is probable IgG4-RD. Even if the patient is younger than 49.5 years old at the first visit, if the serum IgA concentration is < 217 mg/dL, IgG4-RD is still highly likely. If the laboratory findings meet all of the following conditions, IgM < 139.5 mg/dL, IgG $< 1,254$ mg/dL, amylase < 78.5 U/L, and neutrophil count $< 4,414/\mu\text{L}$, there is a strong possibility of IgG4-RD. Alternatively, when the initial node, IgM concentration is ≥ 139.5 mg/dL, the possibility of IgG4-RD is higher when the age at the first visit is ≥ 63 years and the white blood cell count is $< 7,655/\mu\text{L}$, and when the age at the first visit is < 63 years and the IgA concentration is < 151 mg/dL. The ROC curve for this algorithm is shown in Figure 3B. The accuracy, sensitivity, and specificity of the algorithm were 0.823, 0.933, and 0.571, respectively, and the AUC was 0.800. The validation of this algorithm showed that its accuracy, sensitivity, and specificity were 0.732, 0.900, and 0.273, respectively, and the AUC was 0.581 (Figure 3C).

The Gini Impurity is shown in Figure 4A. In this algorithm, the age at the first visit is the most important variable, followed by serum IgM concentration, serum sIL-2R concentration, serum IgE concentration, and serum amylase concentration. The ROC curve for this algorithm is shown in Figure 4B. The accuracy, sensitivity, and specificity of the algorithm were 0.912, 0.990, and 0.714, respectively, and the AUC was

0.981. The validation of this algorithm showed that its accuracy, sensitivity, and specificity were 0.890, 1.000, and 0.591, respectively, and the AUC was 0.919 (Figure 4C).

Discussion

We could firstly in the world construct machine learning algorithm trained to differentiate IgG4-RD from non-IgG4-RD, based on patient characteristics and laboratory findings. Machine learning is one of the research topics in AI. It is a method that learns latent patterns and characteristics from collected data to create models, and it aims to predict and classify new and unknown data.

Recently, the use of AI in the medical field has increased. We have succeeded in building a model for retrospectively identifying abnormalities in diverse types of images [10]. Particularly, systems for detecting colorectal cancer, skin tumors, and cerebral aneurysms, as well as influenza, among others, by AI-based image diagnosis are being developed consecutively. In rheumatology, many results related to treatment support have been reported. Studies have predicted the response to treatment and prognosis of RA by learning clinical markers and genetic analysis data. In 2019, Kim et al. used transcriptome profiling of RA synovium to predict treatment responsiveness from inflammatory signals [11]. Furthermore, Guan et al. have used clinical data and single-nucleotide polymorphism sequence data to predict patient response to antitumor necrosis factor therapy [12]. Since then, many prognostic predictors of patient response to treatment and rehospitalization have been reported for RA, SLE, juvenile idiopathic arthritis, spondyloarthropathy, and osteoarthritis. However, the accuracies of many of these predictors are not yet sufficient [13].

In addition, medical AI can be useful for diagnostic support. We focused our attention on IgG4-RD, a rare group of diseases that have been newly conceptualized in this century. Although awareness is gradually spreading, it has not yet reached general clinicians sufficiently. Diagnosis is based on physical, imaging, serological, and histopathological findings [14]. However, in remote areas, only a few facilities can perform histopathological examinations. IgG4-RD is a homogeneous disorder with certain specific clinical and hematological findings, making it suitable for machine learning. Therefore, we attempted to predict the diagnosis of IgG4-RD in patients needing differentiation by learning only the hematological findings without the imaging and pathological findings.

For this analysis method, the decision tree CART is a learner that selects the variable with the highest impurity and performs branching. The random forest method, which is an ensemble learning method, uses the Gini Impurity calculated from the decision tree CART to perform variable selection. The Gini Impurity is an indicator of the importance of a variable. The random forest method is used as a practical method when a certain number of samples is available. In addition, both functions visualize the reason for the selection, allowing heuristic knowledge acquisition, which is difficult to imagine from the usual way of handling data. In the present study, the diagnosis of IgG4-RD was predicted from the population that needed to be differentiated with high impurity collected from multiple institutions.

In actual clinical practice, the differential diagnosis of IgG4-RD is very important. Depending on the diagnosis, treatment and prognosis may differ greatly. However, since the diseases are uncommon, differential diagnosis can be difficult. Hyper-IgG4emia is not specific to IgG4-RD; it also occurs in eosinophilic diseases such as EGPA, eosinophilic pneumonia, some bronchial asthma, and MCD [15]. In addition, SS and sarcoidosis with the lacrimal and salivary gland swelling, SS with sicca symptoms, and SLE with hypocomplementemia are also important diseases for differentiation.

In addition, about 10%-20% of IgG4-RD cases have associated interstitial pneumonia [16]. For this reason, myositis and vasculitis must also be differentiated as systemic diseases presenting with interstitial pneumonia. Predicting the diagnosis of IgG4-RD in the disease population requiring such differentiation is considered to be of utmost importance. In such cases, when serum IgG4 levels were incorporated, the decision tree showed an AUC of greater than 0.90, and the random forest method showed an AUC of 0.97. When serum IgG4 was not incorporated, the accuracy of differentiation was poor in the decision tree, but the accuracy of differentiation increased to 0.91 AUC in the random forest method.

According to the algorithm of the decision tree, when the serum IgG4 level is known, even if the IgG4 level does not reach 135 mg/dL, which is indicated in the diagnostic criteria, if the level is above the reference value and there is no or a weak inflammatory response, it is likely to be IgG4-RD. This fact is very similar to our actual clinical intuition. However, if the IgG4 level is low, C3 elevation and sIL-2R elevation are not very high, and IgG concentration is almost within the reference range, we do not have the feeling that IgG4-RD is likely. It is expected that IgG4, sIL-2R, and CRP concentrations, which are common to the Gini Impurity at the top, are important for differential diagnosis. In addition, the tree becomes more complicated when the serum IgG4 level is unknown. From the tree nodes and Gini Impurity, age at first diagnosis, IgM, sIL-2R, and amylase concentrations appear to be important factors; IgM may be needed to differentiate MCD and amylase may be to differentiate gland destruction in SS, but this is speculative. The fact that the random forest method was able to predict well even when the IgG4 level was unknown is very interesting and represents a new possibility for AI. These algorithms are not intended for use in the average patient coming in but a tool for a physician to use if trying to distinguish for a patient with these sets of diseases as possible diagnoses without having to resort to more expensive or invasive diagnostic measures.

For AI to be effectively used, a large amount of data is required. As a limitation of this study, the amount of the analyzed data was insufficient even though it was a multicenter study. High-quality data need to be collected to the extent that they can be called "big data". A major task in building machine learning models is to assemble a representative and diverse dataset. Ideally, the model should be trained on data that closely approximate the exact form and quality of data expected at the time of use. Since we focused on the differential diagnosis of IgG4-RD in this study, a bias exists compared to the actual frequency of the disease in clinical practice. In addition, the timing of sample collection differs between IgG4-RD and non-IgG4-RD. The reason for this was that the number of cases of IgG4-RD was small at each institution, and it was not possible to collect a sufficient amount without going back a long time, compared to the control group. To overcome this, using the intractable disease registry that is currently

being constructed in Japan, constructing a database of other diseases and healthy controls that match the frequency of IgG4-RD, building a grand dataset based on this are necessary. In this study, the analysis was conducted using only blood test results. However, it has been highlighted that over-reliance on machine learning models, usually in decision-making and image analysis, may lead to automation bias and reduce the vigilance of physicians against errors [17, 18]. Machine learning applications that directly target patients must pay special attention to this point.

In the near future, when machine learning will be further introduced into lifelong health management [19], we will build a large database that is constantly updated and close to daily clinical practice, considering potential biases. In addition, we want to integrate these clinical and multi-omics data to establish an algorithm that can be applied to the diagnosis of IgG4-RD and all other diseases and that can predict prognosis, such as in drug selection and response to treatment. Based on this study, we hope to use AI to further understand pathological conditions and allow drug discovery.

Conclusions

Our investigations have revealed based on machine learning in a multicenter collaboration, with or without serum IgG4 data, basic patient characteristics and blood test findings alone were sufficient to differentiate IgG4-RD from non-IgG4-RD.

Declarations

Acknowledgments

We would like to express our gratitude to Professor Hiroki Takahashi, Department of Rheumatology and Clinical Immunology, Sapporo Medical University School of Medicine, for permission to conduct this study.

Authors' contributions

MY and MU analyzed the data, and contributed to the writing of the manuscript. MN taught machine learning to MY. RY, SA, IM, TW, AN, TK, and HS collected the sample data in each institution. AK, HY, NY, KK, HY, HH, MK, KT, KF and HT cooperated in the project.

Funding

This work was supported by the JSPS KAKENHI Grant Numbers JP20K08770 and Ministry of Health, Labour and Welfare Research Program on Rare and Intractable Diseases Grant Number JPMH20FC1040.

Available of data and materials

All data generated or analyzed during this study are included in this published article.

Ethics approval and consent to participate

This study was approved by the institutional review boards of the IMSUT Hospital (2019-44-1121), The University of Tokyo Hospital (2019-44-1121), Sapporo Medical University Hospital (322-227), Kanazawa University Hospital (3262-1), Shinshu University Hospital (4652), and Kyoto University Hospital (R2526). Since this study involved medical research on human subjects conducted using existing information, written consent from the patients was not required, and the information was disclosed through the website (opt-out).

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no conflict of interest.

Author details

¹Department of Rheumatology and Allergy, IMSUT Hospital, The Institute of Medical Science, The University of Tokyo. ²Center for Translational Research, IMSUT Hospital, The Institute of Medical Science, The University of Tokyo. ³Department of Otolaryngology, Sapporo Medical University School of Medicine. ⁴Division of Rheumatology, Center for Vaccine and Therapy, IMSUT Hospital, The Institute of Medical Science, The University of Tokyo. ⁵Department of Internal Medicine, Japan Self Defense Sapporo Hospital. ⁶Department of Rheumatology, Kanazawa University Hospital. ⁷Second Department of Internal Medicine, Shinshu University School of Medicine. ⁸Department of Allergy and Rheumatology, Graduate School of Medicine, The University of Tokyo. ⁹Department of Rheumatology and Clinical Immunology, Graduate School of Medicine, Kyoto University. ¹⁰Nagano Prefectural Kiso Hospital.

References

1. Wainberg M, Merico D, DeLong A, Frey BJ. Deep learning in biomedicine. *Nat Biotechnol.* 2018;36:829–36.
2. Stoei BC. Artificial intelligence in detecting early RA. *Semin Arthritis Rheum.* 2019;49:25–8.
3. Hemalatha RJ, Vijaybaskar V, Thamizhvani TR. Automatic localization of anatomical regions in medical ultrasound images of rheumatoid arthritis using deep learning. *Proc Inst Mech Eng H.* 2019;233:657–67.
4. Umehara H, Okazaki K, Masaki Y, Kawano M, Yamamoto M, Saeki T, et al. Comprehensive diagnostic criteria for IgG4-related disease (IgG4-RD), 2011. *Mod Rheumatol.* 2012;22:21–30.
5. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum.* 1997;40:1725.

6. Shiboski CH, Shiboski SC, Seror R, Criswell LA, Labetoulle M, Lietman TM, et al. 2016 American College of Rheumatology/European League Against Rheumatism classification criteria for primary Sjögren's syndrome. *Arthritis Rheumatol.* 2017;69:35–45.
7. Masi AT, Hunder GG, Lie JT, Michel BA, Bloch DA, Calabrese LH, et al. The American College of Rheumatology 1990 criteria for the classification of Churg-Strauss syndrome (allergic granulomatosis and angiitis). *Arthritis Rheum.* 1990;33:1094–100.
8. Lundberg IE, Tjarnlund A, Bottai M, Werth V, Pilkington C, de Visser M, et al. 2017 European League Against Rheumatism/American College of Rheumatology classification criteria for adult and juvenile idiopathic inflammatory myopathies and their major subgroups. *Ann Rheum Dis.* 2017;76:1955–64.
9. Bloch DA, Michel BA, Hunder GG, McShane DJ, Arend WP, Calabrese LH, et al. The American College of Rheumatology 1990 criteria for the classification of vasculitis. *Arthritis Rheum.* 1990;33:1068–73.
10. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. *N Engl J Med.* 2019;380:1347–58.
11. Kim KJ, Kim M, Adamopoulos IE, Tagkopoulos I. Compendium of synovial signatures identifies pathologic characteristics for predicting treatment response in rheumatoid arthritis patients. *Clin Immunol.* 2019;202:1–10.
12. Guan Y, Zhang H, Quang D, Wang Z, Parker SCJ, Pappas DA, et al. Machine learning to predict anti-tumor necrosis factor responses of rheumatoid arthritis patients by integrating clinical and genetic markers. *Arthritis Rheumatol.* 2019;71:1987–96.
13. Jiang M, Li Y, Jiang C, Zhao L, Zhang X, Lipsky PE. Machine learning in rheumatic diseases. *Clin Rev Allergy Immunol.* 2021;60:96–110.
14. Umehara H, Okazaki K, Kawa S, Takahashi H, Goto H, Matsui S, et al. The 2020 revised comprehensive diagnostic (RCD) criteria for IgG4-RD. *Mod Rheumatol.* 2020 Nov 28. doi.org/10.1080/14397595.2020.1859710.
15. Yamamoto M, Tabeya T, Naishiro Y, Yajima H, Ishigami K, Shimizu Y, et al. Value of serum IgG4 in the diagnosis of IgG4-related disease and in differentiation from rheumatic diseases and other diseases. *Mod Rheumatol.* 2012;22:419–25.
16. Yamamoto M, Yajima H, Takahashi H, Yokoyama Y, Ishigami K, Shimizu Y, et al. Everyday clinical practice in IgG4-related dacryoadenitis and/or sialadenitis: results from the SMART database. *Mod Rheumatol.* 2015;25:199–204.
17. Cabitza F, Rasoini R, Gensini GF. Unintended consequences of machine learning in medicine. *JAMA.* 2017;318:517–8.
18. Castelvechi D. Can we open the black box of AI? *Nature.* 2016;5338:20–3.
19. Goecks J, Jalili V, Heiser LM, Gray JW. How machine learning will transform biomedicine. *Cell.* 2020;181:92–101.

Tables

Table 1. Clinical and laboratory items used for machine learning

Clinical findings	Laboratory findings	
Age of the first visit	Rheumatoid factor	Platelet
Sex	Anti-nuclear antibody	LDH
	IgG	AST
	IgG4	ALT
	IgA	γGTP
	IgM	ALP
	IgE	T.bil
	Hemoglobin	D.bil
	CH50	BUN
	C3	Cr
	C4	AMY
	White blood cell	CK
	Neutropil	CRP
	Lymphocyte	soluble IL-2 receptor
Eosinophil		

Table 2. Details of the patients included in the study

	IgG4-RD	Non-IgG4-RD
Male: Female (ratio)	350: 252 (1.4:1)	42: 170 (1:4.0)
Mean age at the first visit (y.o.)	64.11 ± 11.46	54.20 ± 16.09
Mean serum IgG concentration (mg/dL)	2,204.14 ± 1,146.86 (± SD)	1,876.21 ± 1,032.90 (± SD)
Mean serum IgG4 concentration (mg/dL)	666.00 ± 660.46 (± SD)	124.11 ± 397.16 (± SD)
Involved organs or diseases (N)	Dacryoadenitis and sialadenitis (499), Autoimmune pancreatitis (192), Sclerosing cholangitis (51), Kidney disease (107), Respiratory disease (87), Retroperitoneal fibrosis (182)	SS (114), SLE (25), PM/DM (19), MPA (6), EGPA (11), GPA (10), MCD (19), Sarcoidosis (8)
No. of the Involved organs ≥ 2 (%)	350 (58.1)	

IgG4-RD: IgG4-related disease, SS: Sjögren's syndrome, SLE: systemic lupus erythematosus, PM: polymyositis, DM: dermatomyositis, MPA: microscopic polyangiitis, EGPA: eosinophilic granulomatosis with polyangiitis, GPA: granulomatosis with polyangiitis, MCD: multicentric Castleman's disease

Figures

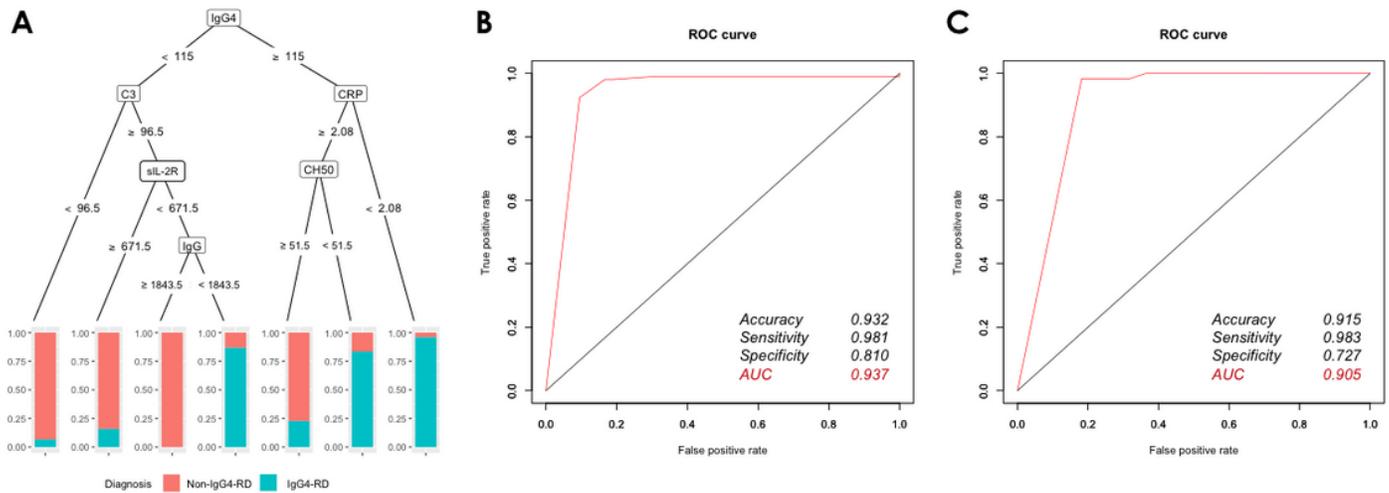


Figure 1

Prediction of IgG4-RD diagnosis in patients with rheumatic diseases requiring differentiation by a CART, when the serum IgG4 level was known A. Decision tree algorithm. The blue color in the figure indicates the predicted percentage of IgG4-RD cases, and the red color indicates percentage of non-IgG4-RD cases. If the serum IgG4 level is ≥ 115 mg/dL, go to the right, and if the CRP is < 2.08 mg/dL, go to the rightmost column, which is probable IgG4-RD. If CRP ≥ 2.08 mg/dL, but CH50 < 51.5 U/mL, the diagnosis of IgG4-RD is more likely. If the laboratory findings meet all of the following conditions, IgG4 < 115 mg/dL, C3 ≥ 96.5 mg/dL, sIL-2R < 671.5 U/mL, and IgG < 1843.5 mg/dL, the likelihood of IgG4-RD increases significantly. B.

ROC curve in the decision tree algorithm. The accuracy, sensitivity, and specificity of the algorithm were 0.932, 0.981, and 0.810, respectively, and the AUC was 0.937. C. ROC curve for the decision tree algorithm (validation). The validation of this algorithm showed that its accuracy, sensitivity, and specificity were 0.915, 0.983, and 0.727, respectively, and the AUC was 0.905.

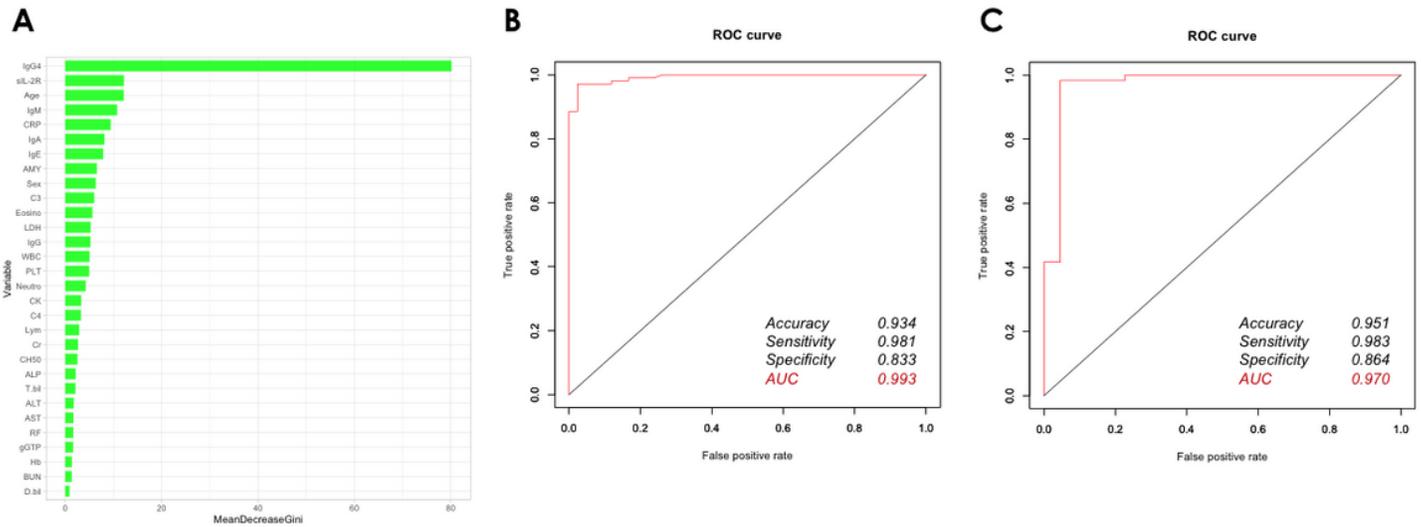


Figure 2

Prediction of IgG4-RD diagnosis in patients with rheumatic diseases requiring differentiation by a random forest, when the serum IgG4 level was known A. Decrease in Gini Impurity. In this algorithm, the serum IgG4 concentration is the most important variable, followed by the serum sIL-2R concentration, age at the first visit, serum IgM concentration, and serum CRP concentration. B. ROC curve for the random forest algorithm. The accuracy, sensitivity, and specificity of the algorithm were 0.934, 0.981, and 0.833, respectively, and the AUC was 0.993. C. ROC curve for the random forest algorithm (validation). The validation of this algorithm showed that its accuracy, sensitivity, and specificity were 0.915, 0.983, and 0.727, respectively, and the AUC was 0.905.

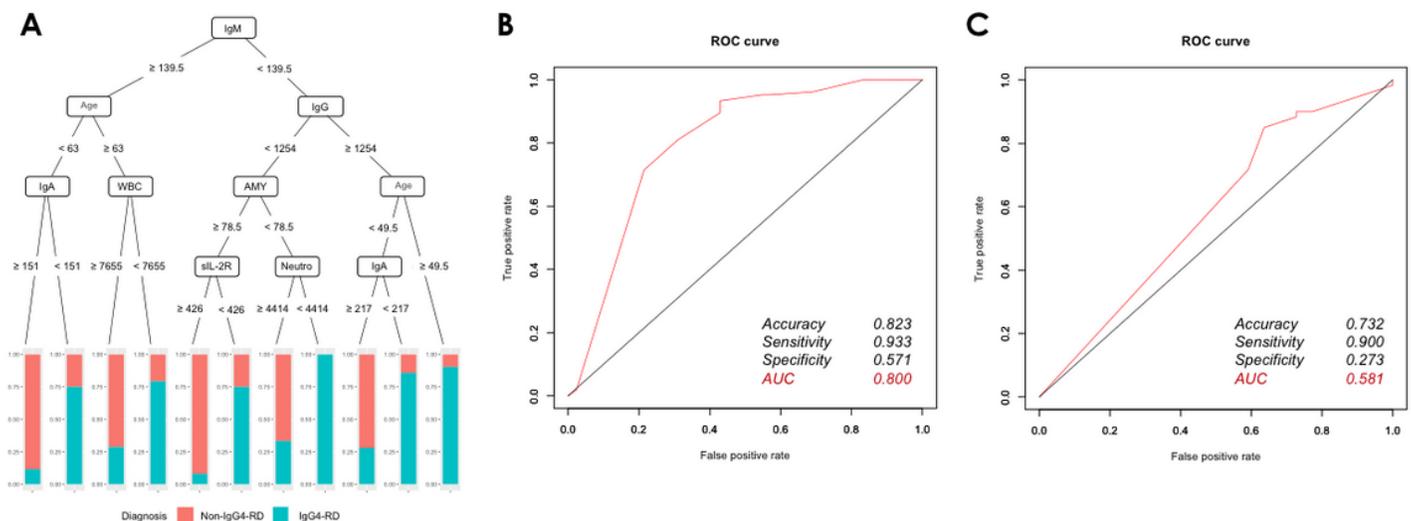


Figure 3

Prediction of IgG4-RD diagnosis in patients with rheumatic diseases requiring differentiation by a CART, when the serum IgG4 level was unknown A. Decision tree algorithm. The blue color in the figure indicates the predicted percentage of IgG4-RD cases, and the red color indicates the percentage of non-IgG4-RD cases. If the serum IgM level is <139.5 mg/dL, go to the right, and if the IgG level is $\geq 1,254$ mg/dL and age at the first visit ≥ 49.5 years, go to the rightmost column, which is probable IgG4-RD. Even if the patient is younger than 49.5 years old at the first visit, if the serum IgA concentration is <217 mg/dL, IgG4-RD is still highly likely. If the laboratory findings meet all of the following conditions, IgM <139.5 mg/dL, IgG $<1,254$ mg/dL, amylase <78.5 U/L, and neutrophil count $<4,414/\mu\text{L}$, there is a strong possibility of IgG4-RD. Alternatively, when the initial node, IgM concentration is ≥ 139.5 mg/dL, the possibility of IgG4-RD is higher when the age at the first visit is ≥ 63 years and the white blood cell count is $< 7,655/\mu\text{L}$, and when the age at the first visit is <63 years and the IgA concentration is < 151 mg/dL. B. ROC curve for the decision tree algorithm. The accuracy, sensitivity, and specificity of the algorithm were 0.823, 0.933, and 0.571, respectively, and the AUC was 0.800. C. ROC curve for the decision tree algorithm (validation). The validation of this algorithm showed that its accuracy, sensitivity, and specificity were 0.732, 0.900, and 0.273, respectively, and the AUC was 0.581.

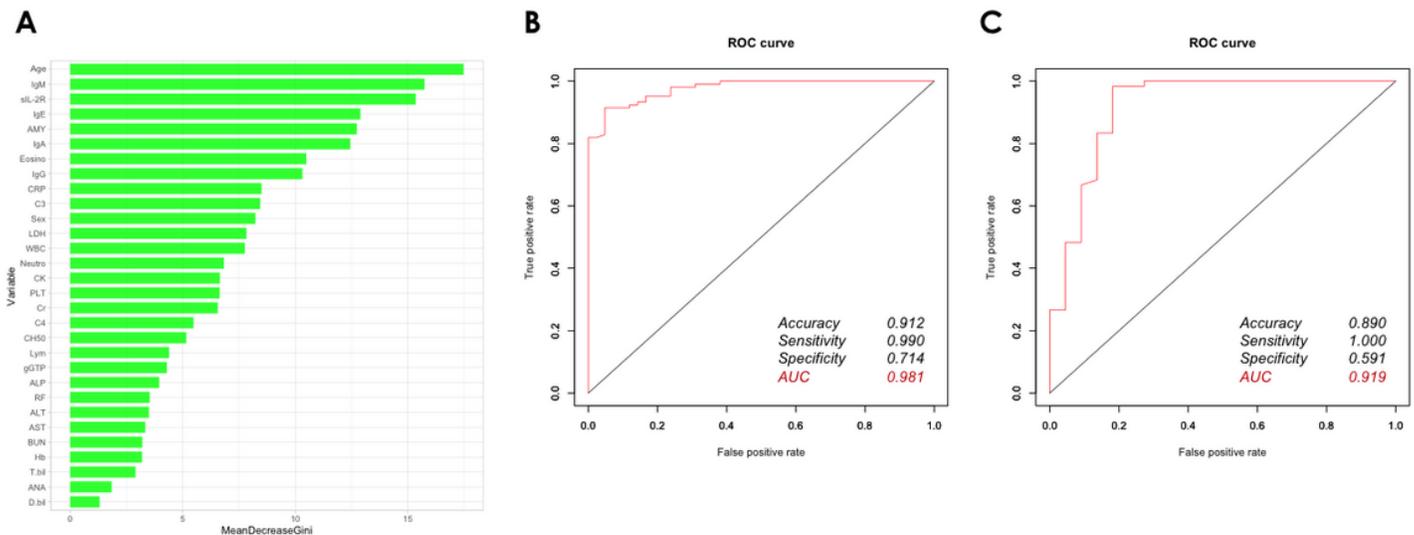


Figure 4

Prediction of IgG4-RD diagnosis in patients with rheumatic diseases requiring differentiation by a random forest, when the serum IgG4 level was unknown A. Decrease in Gini Impurity. In this algorithm, the age at the first visit is the most important variable, followed by serum IgM concentration, serum sIL-2R concentration, serum IgE concentration, and serum amylase concentration. B. ROC curve for the random forest algorithm. The accuracy, sensitivity, and specificity of the algorithm were 0.912, 0.990, and 0.714, respectively, and the AUC was 0.981. C. ROC curve for the random forest algorithm (validation). The validation of this algorithm showed that its accuracy, sensitivity, and specificity were 0.890, 1.000, and 0.591, respectively, and the AUC was 0.919.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Suppl.table1..jpg](#)