

CLINICAL SCIENCE

Clinical phenotypes of IgG4-related disease: an analysis of two international cross-sectional cohorts

Zachary S Wallace,^{1,2,3} Yuqing Zhang,^{1,2,3} Cory A Perugino,^{1,3} Ray Naden,⁴ Hyon K Choi,^{1,2,3} John H Stone,^{1,3} for the ACR/EULAR IgG4-RD Classification Criteria Committee

Handling editor Josef S Smolen

For numbered affiliations see end of article.

Correspondence to

Dr Zachary S Wallace, Rheumatology Unit, Clinical Epidemiology Unit, Division of Rheumatology, Allergy, and Immunology, Massachusetts General Hospital Harvard Medical School, Boston, Massachusetts, USA; zswallace@partners.org

HKC and JHS contributed equally.

This work was previously presented at the European League Against Rheumatism 2018 Annual Meeting (Ann Rheum Dis, volume 77, supplement Suppl, year 2018, page A91) and the American College of Rheumatology Annual Meeting (Arthritis Rheumatol. 2018; 70 (suppl 10)).

Received 17 October 2018
Revised 25 November 2018
Accepted 10 December 2018



© Author(s) (or their employer(s)) 2019. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Wallace ZS, Zhang Y, Perugino CA, et al. *Ann Rheum Dis* Epub ahead of print: [please include Day Month Year]. doi:10.1136/annrheumdis-2018-214603

ABSTRACT

Objective IgG4-related disease (IgG4-RD) is a heterogeneous, multiorgan condition of unclear aetiology that can cause organ failure. Difficulty recognising IgG4-RD contributes to diagnostic delays. We sought to identify key IgG4-RD phenotypes.

Methods We used two cross-sectional studies assembled by an international, multispecialty network of IgG4-RD specialists who submitted 765 cases to derive and replicate phenotypic groups. Phenotype groups of disease manifestations and key covariate distributions across the identified groups were measured using latent class analysis.

Results In the derivation cohort (n=493), we identified four groups with distinct manifestations: Group 1 (31%), Pancreato-Hepato-Biliary disease; Group 2 (24%), Retroperitoneal Fibrosis and/or Aortitis; Group 3 (24%), Head and Neck-Limited disease and Group 4 (22%), classic Mikulicz syndrome with systemic involvement. We replicated the identification of four phenotype groups in the replication cohort. Compared with cases in Groups 1, 2 and 4, respectively, cases in Group 3 were more likely to be female (OR 11.60 (95% CI 5.39 to 24.98), 10.35 (95% CI 4.63 to 23.15) and 9.24 (95% CI 3.53 to 24.20)) and Asian (OR 6.68 (95% CI 2.82 to 15.79), 7.43 (95% CI 2.97 to 18.56) and 6.27 (95% CI 2.27 to 17.29)). Cases in Group 4 had a higher median serum IgG4 concentration (1170 mg/dL) compared with groups 1–3 (316, 178 and 445 mg/dL, respectively, p<0.001).

Conclusion We identified four distinctive IgG4-RD phenotypes according to organ involvement. Being Asian or female may predispose individuals to head and neck-limited disease. These phenotypes serve as a framework for identifying IgG4-RD and studying its aetiology and optimal treatment.

INTRODUCTION

IgG4-related disease (IgG4-RD) is an immune-mediated condition that can affect nearly any organ and often presents with multiorgan involvement.^{1,2} Early recognition and treatment are essential to minimising irreversible organ damage that can result from the disease itself or unnecessary surgical intervention.^{3–6} A challenge to early recognition is the failure to consider the diagnosis, given the manifold patterns of organ involvement with which IgG4-RD can present.³

Clinicians of all specialties need to be able to consider and recognise IgG4-RD in a patient with tumefactive or inflammatory lesions. Previous

Key messages

What is already known about this subject?

- ▶ IgG4-related disease (IgG4-RD) can affect nearly any organ or anatomic site, particularly the salivary glands, pancreato-biliary structures, lymph nodes, lungs, kidneys and retroperitoneum.
- ▶ It was initially described in a Japanese population but has now been described in all racial and ethnic groups.

What does this study add?

- ▶ This study applies a novel cluster analysis method to identify common presentations in the largest multicentre cohort of patients with IgG4-RD assembled by an international collaboration of experts from a variety of specialties.

How might this impact on clinical practice or future developments?

- ▶ We identified four distinctive IgG4-RD phenotypes according to organ involvement which also differed from one another demographically and can be used to frame the approach to diagnosing and studying IgG4-RD.
- ▶ Our findings raise the question of potential racial and/or environmental factors that might influence the risk of certain IgG4-RD manifestations and require further study.

cohort studies have catalogued and described the common and uncommon manifestations of IgG4-RD,^{3,7–9} but patterns of organ involvement remain poorly defined. Defining typical patterns of organ involvement might identify homogenous groups of IgG4-RD, facilitating earlier recognition of the condition. Moreover, these groups may offer a systematic framework that can be used to clarify disease aetiology, identify risk factors and develop personalised treatment strategies.^{10–12}

Given the multiorgan nature of IgG4-RD, the potential number of groups that may be identified is on the order of thousands but most would be neither unique nor useful for clinical application or research.¹³ Latent class analysis (LCA) allows one to identify a parsimonious number of homogenous groups, each composed of individuals who share similar observed characteristics that are distinct

from those defining other groups. To our knowledge, LCA has not been previously applied to group patients with heterogeneous multiorgan diseases such as IgG4-RD.

We used two cohorts assembled by an international, multispecialty network of investigators to perform an LCA to identify distinct phenotypic groups in IgG4-RD.

METHODS

Study population

IgG4-RD specialists from the Americas, Europe and Asia were invited to submit data from cases with either IgG4-RD or a mimicker of IgG4-RD to the two cohorts used to derive and validate the 2018 American College of Rheumatology (ACR) and European League Against Rheumatism (EULAR) Classification Criteria for IgG4-RD.¹⁴ Eighty investigators participated in the ACR/EULAR Classification Criteria Working Group and were invited because of their expertise in IgG4-RD and participation in international symposia. Of the 80 investigators, 52 submitted IgG4-RD cases. The 52 investigators practiced in Japan, China, Australia and 14 countries in the Americas and Europe. There were 25 rheumatologists, 13 gastroenterologists, 3 nephrologists and 11 investigators from other specialties without apparent difference in the distribution of specialties between Asian and non-Asian countries. The larger of the two cohorts included 493 IgG4-RD cases and was used as the primary study population (ie, derivation cohort), whereas the smaller one consisted of 272 IgG4-RD cases and was used to replicate the results (ie, replication cohort). The diagnosis of IgG4-RD was rendered according to the judgement of the participating investigators. All cases were anonymised. This study was approved by the Partners HealthCare Institutional Review Board.

Variables of interest

For each case, investigators submitted details including age (at symptom/disease onset and diagnosis), sex, race/ethnicity, organ involvement and serum IgG4 concentrations. Time to diagnosis was calculated by subtracting the age at symptom onset from age at diagnosis, reported in years. In the derivation cohort, investigators reported each case's serum IgG4 concentration, the associated unit of measurement and the upper limit of normal (ULN) in the laboratory in which the assay was performed. All serum IgG4 concentrations were converted to milligram per decilitre (mg/dL). Based on the reported ULN, we created four serum IgG4 concentration categories. In the replication cohort, only the serum IgG4 concentration category was reported. Based on reported race/ethnicity, we dichotomised cases into either Asian or non-Asian. We chose to dichotomise race by Asian or non-Asian because of differences in organ distribution, serum IgG4 concentration elevations and other factors observed when comparing Asian and non-Asian cohorts.¹⁵ Patients of South Asian descent (eg, India, Pakistan; n=14), all of whom resided in North America or Europe, were grouped with non-Asian cases.

Statistical analysis

We described the characteristics of study patients overall and according to sex and race (Asian vs non-Asian). For continuous covariates, summary statistics were reported as mean and SD or median and IQR, where appropriate. Proportions were compared using the Chi square test and continuous variables were compared using the Student's t test or the Wilcoxon rank-sum test, as appropriate.

We performed a LCA using SAS procedure PROC LCA to classify subjects into mutually exclusive groups based on their

clinical manifestations of IgG4-RD.¹⁶ Twenty-eight potential manifestations (online supplementary table 1), each representing different organs or anatomic sites (eg, submandibular gland, pancreas), were used as categorical variables to identify groups (eg, latent classes).

We began by fitting models with potential solutions ranging from 2 to 5 groups in the primary study cohort. Of the models that ranged in size from 2 to 5 groups, we chose the best fit model based on the lowest Akaike information criteria (AIC) and adjusted Bayesian information criterion (BIC). When five or more groups were fit in the model, the AIC and BIC continued to increase. The probability of group membership for each patient was obtained from the LCA model and each patient was assigned to the group in which he or she had the highest probability of membership.¹⁶

We then incorporated a set of key covariates (ie, sex, race, age at diagnosis, time to diagnosis and serum IgG4 concentration) into the model with the best fit. These covariates were chosen based on a priori knowledge of IgG4-RD and prior studies.³ The distribution of organ involvement in each group is reported as a probability. We labelled each group with a descriptive term based on the common manifestation(s) present in each group and the distribution of organ involvement within each group. We examined the relation of each covariate to the odds of belonging to a specific group using multivariable logistic regression. We performed a sensitivity analysis in which we only included cases that fulfilled the proposed 2018 ACR/EULAR Classification Criteria.¹⁴ We applied the same approach described above in our replication cohort to determine the optimal number of groups, manifestations characteristic of each group and associations between covariates and group membership. We used SAS 9.3 (SAS Institute, Cary, North Carolina, USA) for all analyses. A two-sided $p < 0.05$ was assumed to be significant for all analyses.

RESULTS

Of the 493 cases in the primary cohort, 322 (65.3%) were male and 285 (57.8%) were non-Asian with 198 (40.2%) Caucasians (table 1). Rheumatologists submitted 193 (39%) of these cases, 90 from Asian countries and 103 from non-Asian countries. The mean (SD) ages at symptom onset and diagnosis were 57.7 (14.5) years and 59.5 (14.0) years, respectively, resulting in a mean (SD) time to diagnosis of 1.8 (3.4) years. The characteristics of the replication cohort (n=272) were similar (online supplementary table 3).

An elevated serum IgG4 concentration was reported in 388 (78.7%) cases (table 1). The mean (SD) number of organs affected in each case was 2.9 (1.8). The most commonly affected organs were in the pancreato-hepato-biliary system (235, 47.7%), followed by involvement of the salivary glands (186, 37.7%). At least one biopsy was performed in 425 (85.4%) cases, but the classically described histological feature of storiform fibrosis was reported in only 195 (39.6%) cases. Male patients were older (mean (SD)) at the time of symptom onset (59.9 (14.3) vs 53.4 (14.0) years; $p < 0.001$) and diagnosis (61.7 (13.8) vs 55.4 (13.5) years; $p < 0.001$) than female patients and were less likely to have head and neck disease (46.9% vs 65.5%, $p < 0.001$) but more likely to have retroperitoneal fibrosis (RPF) (19.3% vs 9.4%, $p = 0.004$) (table 1).

Asian patients were older (mean (SD)) at symptom onset (61.2 (13.2) years vs 55.1 (14.9) years, $p < 0.001$) and diagnosis (62.6 (12.8) years vs 57.2 (14.4) years, $p < 0.001$), had a higher baseline median (IQR) serum IgG4 concentration (666 (321-1,230) mg/dL vs 240.5 (100-505) mg/dL, $p < 0.001$) and more often had

Table 1 Demographic features overall and by sex and race (derivation cohort)

	Sex			Race	
	All (n=493)	Male (n=322)	Female (n=171)	Asian (n=208)	Non-Asian (n=285)
Age at symptom onset (years, mean, SD)	57.7 (14.5)	59.9 (14.3)	53.4 (14.0)	61.2 (13.2)	55.1 (14.9)
Age at diagnosis	59.5 (14.0)	61.7 (13.8)	55.4 (13.5)	62.6 (12.8)	57.2 (14.4)
Time to diagnosis	1.8 (3.4)	1.7 (3.6)	2.0 (2.9)	1.4 (2.7)	2.2 (3.7)
Male (n,%)	322 (65.3%)	–	–	126 (60.6%)	196 (68.7%)
Race					
Caucasian	198 (40.2%)	144 (44.7%)	54 (31.6%)	–	198 (69.5%)
Asian*	208 (42.2%)	126 (39.1%)	82 (47.9%)	208 (42%)	–
Asian, not otherwise specified	48 (9.7%)	34 (10.6%)	14 (8.2%)	48 (9.7%)	–
Japanese	105 (21.3%)	61 (18.9%)	44 (25.7%)	105 (21.3%)	–
Chinese	53 (10.8%)	29 (9.0%)	24 (14.0%)	53 (10.8%)	–
Latino/Hispanic	58 (11.8%)	32 (9.9%)	26 (15.2%)	–	58 (20.4%)
South Asian	14 (2.8%)	10 (1.1%)	4 (2.3%)	–	14 (4.9%)
Black	9 (1.8%)	8 (2.5%)	5 (2.9%)	–	9 (3.2%)
Other	6 (1.2%)	2 (0.6%)	0 (0%)	–	6 (2.1%)
IgG4 concentration (mg/dL, median, IQR)	342.5 (160–870)	382 (177–870)	302 (124–866)	666 (321–1230)	240.5 (100–505)
Not checked	15 (3.0%)	13 (4.0%)	2 (1.2%)	0 (0%)	15 (5.3%)
Normal	90 (18.3%)	49 (15.2%)	41 (24.0%)	17 (8.2%)	73 (25.6%)
>Normal < 2×ULN	83 (16.8%)	54 (16.8%)	29 (17.0%)	24 (11.5%)	59 (20.7%)
>2×Normal <5×ULN	149 (30.2%)	98 (30.4%)	51 (29.8%)	67 (32.2%)	82 (28.8%)
>5×ULN	156 (31.6%)	108 (33.5%)	48 (28.1%)	100 (48.1%)	56 (19.6%)
Biopsy performed	421 (85.4%)	267 (82.9%)	154 (90.1%)	188 (90.4%)	233 (81.8%)
Obliterative phlebitis	91 (18.5%)	55 (17.1%)	36 (21.1%)	19 (9.1%)	72 (25.3%)
Storiform fibrosis	195 (39.6%)	126 (39.1%)	69 (40.4%)	57 (27.4%)	138 (48.4%)
Number of organs affected (mean, SD)	2.9 (1.8)	2.9 (1.9)	2.8 (1.6)	2.9 (1.5)	2.9 (1.9)
Single organ	120 (24.3%)	75 (23.3%)	45 (26.3%)	33 (15.9%)	87 (30.5%)
Multiorgan (≥2 organs)	373 (75.6%)	247 (76.7%)	126 (73.7%)	175 (84.1%)	198 (69.5%)
Head and Neck	263 (53.3%)	151 (46.9%)	112 (65.5%)	141 (67.8%)	122 (42.8%)
Salivary gland	186 (37.7%)	110 (34.2%)	76 (44.4%)	109 (52%)	77 (27.0%)
Orbital	32 (6.5%)	15 (4.7%)	17 (9.9%)	8 (3.9%)	24 (8.4%)
Extraocular muscle	20 (4.1%)	6 (1.9%)	14 (8.2%)	5 (2.4%)	15 (5.3%)
Lacrimal gland	128 (26.0%)	61 (18.9%)	67 (39.2%)	80 (38.5%)	48 (16.8%)
Pancreato-hepatobiliary	235 (47.7%)	178 (55.3%)	57 (33.3%)	88 (42.3%)	147 (51.6%)
Pulmonary	70 (14.2%)	51 (15.8%)	19 (11.1%)	30 (14.4%)	40 (14.0%)
Aorta	51 (10.3%)	38 (11.8%)	13 (7.6%)	20 (9.6%)	31 (11.0%)
Retroperitoneum	78 (15.8%)	62 (19.3%)	16 (9.4%)	28 (13.5%)	50 (17.5%)
Renal	77 (15.6%)	58 (18.0%)	19 (11.1%)	34 (16.3%)	43 (15.1%)

*Two patients (0.4%) were Korean, both were male.
ULN, upper limit of normal.

head and neck disease (67.8% vs 42.8%, $p < 0.001$) (table 1). Of note, there was a shorter time to diagnosis (mean (SD)) among Asian patients compared with non-Asian patients (1.4 (2.7) vs 2.2 (3.7) years, $p = 0.01$). Similar differences were observed in the replication cohort according to sex and race (online supplementary table 2).

Determination of phenotype groups

After fitting candidate cluster models (2–5 groups), we found that the four-group model had the best fit statistics and aligned with clinical experience (online supplementary table 3). The four groups were distinguished from one another by the distribution of organ involvement; the average posterior probability of group membership ranged from 90% to 93%, indicating mutually exclusive group assignment (table 2 and figure 1).

Group 1 (n=149, 31%) was characterised by pancreato-hepatobiliary disease, whereas Group 2 (n=114, 24%) was characterised by RPF and/or aortitis involvement. Group 3 (n=115, 24%)

was characterised by disease generally limited to the head and neck structures in a pattern of incomplete Mikulicz syndrome. The probability of parotid gland involvement was only 22% in Group 3. Group 4 (n=100, 22%) was characterised by head and neck disease in a pattern more consistent with Mikulicz syndrome along with extraglandular, systemic involvement. In contrast to Group 3, the probability of parotid gland involvement was 49% in Group 4. In addition, there was a 22% probability of orbital disease in Group 3 but <1% probability in Group 4. The mean (SD) number of organs affected was significantly larger in Group 4 (5.2 (1.9)) than in others (2.1 (1.1), 2.1 (1.2) and 2.7 (1.4) for Groups 1–3, respectively, $p < 0.001$).

The distribution of key covariates according to group membership and associations between them are shown in (tables 3 and 4). Compared with Group 3 (Head and Neck-Limited) in which 76% were female, Group 1 (Pancreato-Hepato-Biliary), Group 2 (RPF/Aorta) and Group 4 (Mikulicz and Systemic) had a low proportion of female patients (21%, 26% and 22%, respectively).

Table 2 Phenotypic groups of IgG4-RD (derivation cohort)*

Variables used to identify group†	Group 1 'Pancreato-Hepato-Biliary' (%)	Group 2 'Retroperitoneum and Aorta' (%)	Group 3 'Head and Neck-Limited' (%)	Group 4 'Mikulicz and Systemic' (%)	P value
Pancreas	87	12	15	46	<0.001
Liver	13	1	2	5	<0.001
Biliary	55	<1	<1	27	<0.001
Orbital	<1	3	22	<1	<0.001
Extraocular muscle	<1	1	13	4	<0.001
Sinusitis	3	<1	17	16	<0.001
Parotid gland	2	1	22	49	<0.001
Submandibular gland	15	5	50	77	<0.001
Lacrimal gland	3	3	60	48	<0.001
Renal	11	13	5	36	<0.001
Lung	2	15	7	39	<0.001
Lymph node	15	25	29	67	<0.001
Prostate	1	<1	<1	14	<0.001
Thoracic aorta	1	10	1	3	<0.001
Abdominal aorta	3	22	<1	13	<0.001
Retroperitoneum	4	53	2	8	<0.001
Proportion of cohort	31	24	24	22	
Average probability (Mean)	93	92	90	90	

*Only cases (n=478) with complete covariate data are included in the latent class analysis with covariates.

†All percentages are probabilities that the manifestation occurs in that group, conditional on latent class membership (eg, totals will not equal 100%). IgG4-RD, IgG4-related disease.

The adjusted ORs for female patients being in Group 3 vs Groups 1, 2 and 4 were 11.6 (95% CI 5.4 to 25.0), 10.4 (95% CI 4.6 to 23.2) and 9.2 (95% CI 3.5 to 24.2), respectively. The distribution of race/ethnicity also differed across groups. Proportions of Asian patients in Groups 3 (67%) and 4 (52%) were much higher than in Groups 1 (37%) and 2 (25%). The adjusted ORs for Asians being in Group 3 vs Groups 1, 2 and 4 were 6.7 (95% CI 2.8 to 15.8), 7.4 (95% CI 3.0 to 18.6) and 6.3 (95% CI 2.3 to 17.3), respectively (table 4). Mean age at diagnosis was younger in Groups 3 than in the other three clusters (p<0.001)

but patients in Group 3 tended to have a longer time to diagnosis than those in Groups 1 and 2 (p<0.001), respectively.

Serum IgG4 concentrations also differed across the four groups. Patients in Group 4 (Mikulicz and Systemic) had the highest serum IgG4 concentrations (median (IQR) 1170 (520-2178) mg/dL), followed by Group 3 (Head and Neck-Limited; 445 (183-888) mg/dL) and Group 1 (Pancreato-Hepato-Biliary; 316 (147-622) mg/dL); those in Group 2 (RPF/Aorta) had the lowest serum IgG4 concentrations (178 (63-322) mg/dL). Compared with Group 2, for every 100 mg/dL increase in the

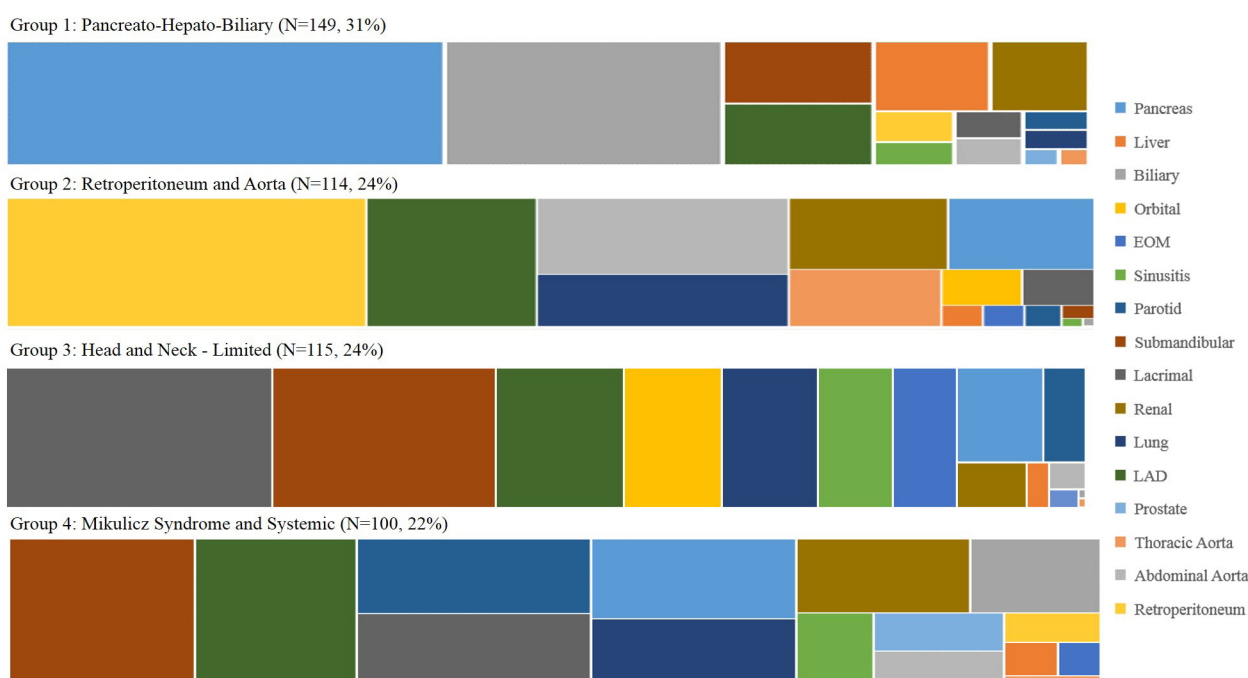


Figure 1 Distribution of organ involvement in each group (% of overall cohort). EOM, extraocular muscle; Lad, lymphadenopathy.

Table 3 Demographics and key covariates according to IgG4-RD phenotype groups (derivation cohort)

Covariate	Group 1 'Pancreato-Hepato-Biliary'	Group 2 'Retroperitoneum and Aorta'	Group 3 'Head and Neck-Limited'	Group 4 'Mikulicz and Systemic'
Female (%)	21%	25%	76%	22%
Asian (%)	37%	25%	67%	52%
Age at diagnosis (year, mean, SD)	63 (13)	58 (16)	55 (13)	63 (13)
Time to diagnosis (year, mean, SD)	0.9 (1.8)	1.8 (4.0)	2.3 (3.4)	2.0 (3.6)
Serum IgG4 concentration (mg/dL, median, IQR)	316 (147–622)	178 (63–322)	445 (183–888)	1170 (520–2178)

serum IgG4 concentration, the adjusted OR of being in Group 3 was 1.2 (95% CI 1.1 to 1.4). In contrast, compared with Group 4, the adjusted OR of being in Group 3 was 0.9 (95% CI 0.84 to 0.96) for every 100 mg/dL increase in the serum IgG4 concentration.

In a sensitivity analysis, that included only cases that fulfilled the proposed 2018 ACR/EULAR Classification Criteria (86% in Group 1, 77% in Group 2, 84% in Group 3 and 84% in Group 4), our results did not change materially.¹⁴

Using the replication cohort (online supplementary table 2), we identified four phenotypic groups with the same organ distribution characteristics (online supplementary table 5) as those identified using the derivation cohort. As in the derivation cohort, female and Asian patients tended to be in the groups characterised by head and neck disease; the Mikulicz and Systemic Group had the highest serum IgG4 concentrations (online supplementary tables 6 and 7).

DISCUSSION

Multiorgan diseases such as IgG4-RD pose challenges in aetiological understanding, in part because of their heterogeneous manifestations. Using an unbiased approach—latent class analysis (LCA)—that identifies groups sharing common features, we characterised four common phenotypes of IgG4-RD based on organ involvement patterns. The groups identified in this manner also differed significantly from one another by age, sex, race, serum IgG4 concentration and time to diagnosis.

Our results suggest that cases of Asian descent are particularly predisposed to developing IgG4-RD complications in the head and neck region and have a predilection for disease limited to this region. In contrast, non-Asian cases—predominantly Caucasians—have a greater predilection for pancreato-hepatobiliary disease and/or retroperitoneal and aorta disease compared with Asian patients. It is noteworthy that the first recognition of pancreato-hepatobiliary disease associated with IgG4-RD occurred in Japan.¹⁷ Our findings are the first to imply that the distribution of organs affected by IgG4-RD differs between Asian and non-Asian cases. Differences observed in variables such as race, age and serum IgG4 concentrations across groups suggest

that genetic or environmental risk factors may differ across groups and, perhaps, among Asian and non-Asian patients. Given the similar distribution of subspecialists among investigators in this study practicing in Asian and non-Asian countries, the observed differences are unlikely to be the result of detection or selection biases.

Our observation of race as a potential risk factor for certain IgG4-RD manifestations adds to a growing body of literature describing potential IgG4-RD risk factors, including tobacco and asbestos exposure,^{18 19} atopic disease^{20–22} and a history of malignancy,²³ which may also differ across clusters. For instance, a prior study found that tobacco and asbestos exposure were risk factors for 'idiopathic' RPF, a common IgG4-RD manifestation which tended to be present in Group 2.¹⁸ Moreover, our group has identified associations between atopic phenotypes and head and neck manifestations²² as well as an increased flare risk.²¹ Future studies should evaluate whether certain phenotypes are more likely to have atopic phenotypes or other comorbidities.

This multicentre study confirms the findings of earlier, single-centre studies that cases with multiorgan disease (eg, Group 4) often have higher serum IgG4 concentrations,^{3 7} whereas RPF cases (eg, Group 2) generally have lower concentrations.³ Although previous studies have suggested that female patients are more likely to have head and neck disease, our findings specify that females tend to have disease limited to the head and neck rather than systemic disease with head and neck manifestations.⁷

Time to diagnosis varied across the groups but tended to be long, underscoring the persistent challenges in diagnosing IgG4-RD. The interval between symptom onset and receipt of the IgG4-RD diagnosis was nearly 2 years overall but varied from approximately 1 year in Group 1 (Pancreato-Hepato-Biliary) to over 2 years in Groups 3 (Head and Neck-Limited) and 4 (Mikulicz and Systemic). Facilitating early diagnosis is critical: the observed association between a group with greater number of organs affected and a longer diagnostic delay suggests that diagnostic delays permit the accrual of additional organ involvement, heightening the risk of permanent organ damage (eg, pancreatic insufficiency, renal failure).^{3–6}

Table 4 Multivariable logistic regression (adjusted OR, 95% CI) evaluating the odds of belonging to the Head and Neck-Limited group based on selected covariates (derivation cohort)

Covariate	'Head and Neck-Limited' versus 'Pancreato-Hepato-Biliary'	'Head and Neck-Limited' versus 'Retroperitoneum and Aorta'	'Head and Neck-Limited' versus 'Mikulicz and Systemic'
Female	11.60 (5.39 to 24.98)	10.35 (4.63 to 23.15)	9.24 (3.53 to 24.20)
Asian	6.68 (2.82 to 15.79)	7.43 (2.97 to 18.56)	6.27 (2.27 to 17.29)
Age at diagnosis (per 5 years)	0.78 (0.67 to 0.89)	0.85 (0.74 to 0.98)	0.80 (0.68 to 0.95)
Time to diagnosis (per year)	1.41 (1.19 to 1.66)	1.14 (1.04 to 1.26)	1.12 (0.99 to 1.25)
Serum IgG4 concentration (per 100 mg/dL)	1.03 (0.96 to 1.11)	1.20 (1.05 to 1.36)	0.90 (0.84 to 0.96)

IgG4-RD, IgG4-related disease.

The longer time to diagnoses observed in some groups is likely multifactorial in nature. First, important knowledge deficits regarding the existence of this condition—recognised only within the last 15 years—persist among medical practitioners. Second, for many manifestations, existing criteria that define IgG4-RD are weighted heavily to serum IgG4 concentrations and the presence of certain pathology findings.²⁴ The forthcoming ACR/EULAR Classification Criteria emphasise the importance of clinical and radiological features as well as serological and pathological findings. These criteria will facilitate the identification of patients appropriate for study but will also orient diagnostic thinking in a more expansive way.

Our findings suggest other steps that may optimise IgG4-RD patient care. The association between higher IgG4 concentrations and more extensive disease (eg, Group 4) suggests that in cases with high IgG4 concentrations, one should carefully review the case's history, perform a full physical examination, and consider advanced imaging (eg, CT) to identify other organ involvement when disease seems isolated to a single organ.

Our study has several strengths. We used an unbiased approach to analyse two of the largest IgG4-RD cohorts representing a wide disease spectrum assembled through an international collaboration of experts from a variety of specialties. Investigators were encouraged to include all cases that they considered to have IgG4-RD, regardless of whether the case fulfilled previously published diagnostic or classification criteria. Moreover, we replicated the four phenotypes and other observations identified from one cohort in the second cohort, indicating the robustness of our results.

Despite these strengths, our study has certain limitations. First, some manifestations known to occur infrequently (eg, hypophysitis) were rarely reported in this cohort, limiting our ability to describe their distribution across the groups. Second, this was a cross-sectional study in which cases were reported for the purposes of validating IgG4-RD classification criteria. Thus, the assessment of cases was not standardised such that the use of advanced imaging and certain laboratory tests likely varied by investigator. Moreover, we cannot rule out the potential effect of selection bias on our findings regarding race. However, there was no difference in the distribution of subspecialists between Asian and non-Asian countries, we encouraged investigators to submit all cases which they considered to be IgG4-RD regardless of whether they had fulfilled previous criteria, our findings remained unchanged when we limited our analysis to cases fulfilling the proposed ACR/EULAR Classification Criteria, and there was no difference in the distribution of diagnostic confidence between investigators from Asian and non-Asian countries. Third, we had few cases of Asian descent evaluated in non-Asian countries so we are unable to meaningfully discern the relative impact of environment as opposed to race on our findings. Fourth, participation by investigators with expertise in IgG4-RD at academic centres may limit the generalisability of our findings. However, this is a relatively rare disease that is often diagnosed and/or managed by clinicians like those who participated in this study. We therefore suspect that we captured the wide range of disease likely to be encountered by physicians in a variety of settings. Fifth, while we replicated the phenotypes observed in our derivation cohort, the small sample size of our replication cohort limited our ability to detect some of the differences observed in multivariable analyses of the derivation cohort.

In conclusion, we have described four distinctive phenotypes of IgG4-RD according to the distribution of organ involvement. In addition to differences in organ involvement, these groups

also differed by race, age, time to diagnosis, sex and serum IgG4 concentration. Differences in genetic and environmental risk factors associated with each group may explain these observations. These phenotypes may be used by clinicians to improve recognition of IgG4-RD. These findings form the basis of further investigations designed to understand these factors in greater detail.

Author affiliations

¹Rheumatology Unit, Division of Rheumatology, Allergy, and Immunology, Massachusetts General Hospital, Boston, Massachusetts, USA

²Clinical Epidemiology Unit, Division of Rheumatology, Allergy, and Immunology, Massachusetts General Hospital, Boston, Massachusetts, USA

³Harvard Medical School, Boston, Massachusetts, USA

⁴Department of Medicine, McMaster University, Hamilton, Ontario, Canada

Correction notice This article has been corrected since it published Online First. The collaborators statement has been corrected.

Acknowledgements The IgG4-Related Disease Classification Criteria Working Group of the American College of Rheumatology and European League Against Rheumatism (see list at the end of the manuscript).

Collaborators Takashi Akamizu; Mitsuhiro Akiyama; Adrian Bateman; Daniel Blockmans; Pilar Brito-Zeron; Corrado Campochiaro; Mollie Carruthers; Suresh Chari; Tsutomu Chiba; Andreu Fernandez Codina; Lynn Cornell; Emma Culver; Emanuel Della-Torre; Vikram Deshpande; Jean-Francois Dicaire; Lingli Dong; Mikael Ebbo; Judith A Ferry; George Fragkouli; Fabian Frost; Luca Frulloni; Phil A Hart; Gabriela Hernandez-Molina; Dai Inoue; Karuna Keat; Terumi Kamisawa; Shigeyuki Kawa; Mitsuhiro Kawano; Arezou Khosroshahi; Hiroshi Kobayashi; Yuzo Kodama; Satoshi Kubo; Kensuke Kubota; Marco Lanzillotta; Markus Lerch; Yanying Liu; Matthias Löhr; Chiara Marvisi; Ferran Martinez-Valle; Eduardo Martin-Nares; Yasufumi Masaki; Shoko Matsui; Ichiro Mizushima; Seiji Nakamura; Jan Nordeide; Kenji Notohara; Kazuichi Okazaki; Sergio Pairea; Jovan Popovic; Manel Ramos-Casals; James Rosenbaum; Jay Ryu; Yasuhiro Sato; Amita Sharma; Takako Saeki; Hiroshi Sekiguchi; Nicolas Schleinitz; Evgeniya V Sokol; James R Stone; Hiroki Takahashi; Naoki Takahashi; Masayuki Takahira; Yoshiya Tanaka; Hisanori Umehara; Augusto Vaglio; Alejandra Villamil; Yoko Wada; George Webster; Kazunori Yamada; Motohisa Yamamoto; Joanne Yi; Giuseppe Zamboni; Yoh Zen; Wen Zhang

Contributors All named authors contributed to planning, conduct and reporting of the work. Members of the IgG4-Related Disease Classification Criteria Working Group contributed to conduct of the study because they contributed detailed information regarding cases.

Funding ZSW received grant support through a Scientist Development Award from the Rheumatology Research Foundation and from the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS/NIH; Loan Repayment Award and K23 AR073334). JHS received funding for development of the IgG4-RD Classification Criteria from the American College of Rheumatology and the European League Against Rheumatism.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval This study was approved by the Partners HealthCare Institutional Review Board.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Investigators interested in collaboration or using the data from the IgG4-Related Disease Classification Criteria Working Group can contact Dr John H Stone (jstone@mgh.harvard.edu). Requests will be considered on a case-by-case basis.

REFERENCES

- Stone JH, Zen Y, Deshpande V. IgG4-related disease. *N Engl J Med* 2012;366:539–51.
- Kamisawa T, Zen Y, Pillai S, et al. IgG4-related disease. *Lancet* 2015;385:1460–71.
- Wallace ZS, Deshpande V, Mattoo H, et al. IgG4-related disease: clinical and laboratory features in one hundred twenty-five patients. *Arthritis Rheumatol* 2015;67:2466–75.
- Hirano K, Tada M, Isayama H, et al. Long-term prognosis of autoimmune pancreatitis with and without corticosteroid treatment. *Gut* 2007;56:1719–24.
- Hirano K, Isogawa A, Tada M, et al. Long-term prognosis of autoimmune pancreatitis in terms of glucose tolerance. *Pancreas* 2012;41:1–95.
- Shimizu Y, Yamamoto M, Naishiro Y, et al. Necessity of early intervention for IgG4-related disease—delayed treatment induces fibrosis progression. *Rheumatology* 2013;52:679–83.
- Inoue D, Yoshida K, Yoneda N, et al. IgG4-related disease: dataset of 235 consecutive patients. *Medicine* 2015;94:e680.

- 8 Sekiguchi H, Horie R, Kanai M, *et al.* IgG4-related disease: retrospective analysis of one hundred sixty-six patients. *Arthritis Rheumatol* 2016;68:2290–9.
- 9 Hart PA, Kamisawa T, Brugge WR, *et al.* Long-term outcomes of autoimmune pancreatitis: a multicentre, international analysis. *Gut* 2013;62:1771–6.
- 10 Roduit C, Frei R, Depner M, *et al.* Phenotypes of atopic dermatitis depending on the timing of onset and progression in childhood. *JAMA Pediatr* 2017;171:655–62.
- 11 Kittelson AJ, Stevens-Lapsley JE, Schmiede SJ. Determination of pain phenotypes in knee osteoarthritis: a latent class analysis using data from the osteoarthritis initiative. *Arthritis Care Res* 2016;68:612–20.
- 12 Perugino CA, Mattoo H, Mahajan VS, *et al.* Emerging treatment models in rheumatology: IgG4-related disease: insights into human immunology and targeted therapies. *Arthritis Rheumatol* 2017;69:1722–32.
- 13 Lanza ST, Rhoades BL. Latent class analysis: an alternative perspective on subgroup analysis in prevention and treatment. *Prev Sci* 2013;14:157–68.
- 14 Wallace ZS, Naden RP, Choi H. The 2018 ACR/EULAR Classification Criteria for IgG4-Related Disease. *Submitted for simultaneous publication in Arthritis & Rheumatology and Annals of the Rheumatic Diseases* 2018.
- 15 Qi R, Chen LYC, Park S, *et al.* Utility of serum IgG4 levels in a multiethnic population. *Am J Med Sci* 2018;355:61–6.
- 16 Lanza ST, Collins LM, Lemmon DR, *et al.* PROC LCA: a SAS procedure for latent class analysis. *Struct Equ Modeling* 2007;14:671–94.
- 17 Aithal GP, Breslin NP, Gumustop B. High serum IgG4 concentrations in patients with sclerosing pancreatitis. *N Engl J Med* 2001;345:147–8.
- 18 Goldoni M, Bonini S, Urban ML, *et al.* Asbestos and smoking as risk factors for idiopathic retroperitoneal fibrosis: a case-control study. *Ann Intern Med* 2014;161:181–8.
- 19 Maire F, Rebours V, Vullierme MP, *et al.* Does tobacco influence the natural history of autoimmune pancreatitis? *Pancreatol* 2014;14:284–8.
- 20 Della Torre E, Mattoo H, Mahajan VS, *et al.* Prevalence of atopy, eosinophilia, and IgE elevation in IgG4-related disease. *Allergy* 2014;69:269–72.
- 21 Wallace ZS, Mattoo H, Mahajan VS, *et al.* Predictors of disease relapse in IgG4-related disease following rituximab. *Rheumatology* 2016;55:1000–8.
- 22 Sanders S, Della Torre E, Perugino CA, *et al.* Salivary Gland Disease in IgG4-Related Disease Is Associated with Allergic Histories [abstract]. *Arthritis Rheumatol* 2018;70(suppl 10).
- 23 Wallace ZS, Wallace CJ, Lu N, *et al.* Association of IgG4-related disease with history of malignancy. *Arthritis Rheumatol* 2016;68:2283–9.
- 24 Okazaki K, Umehara H. Are Classification criteria for IgG4-RD now possible? the concept of igg4-related disease and proposal of comprehensive diagnostic criteria in Japan. *Int J Rheumatol* 2012;2012:1–9.