RHEUMATOLOGY

Rheumatology 2021;60:217–223
doi:10.1093/rheumatology/keaa247
Advance Access publication 13 July 2020

Original article

Performance of the 2019 ACR/EULAR classification criteria for IgG4-related disease and clinical phenotypes in a Spanish multicentre registry (REERIGG4)

Andreu Fernández-Codina[1,2], Blanca Pinilla[3], Iago Pinal-Fernández[4], Iago Carballo[5], Carlos Feijoo-Massó[6], María Toledano-Macías[3], Borja de Miguel-Campo[7], Eva Fonseca-Aizpuru[8], Luis Sáez-Comet[9], Miguel López-Dupla[10], José Hernández-Rodríguez[11] and Fernando Martínez-Valle[1]

Abstract

Objectives. Several IgG4-related disease (IgG4-RD) phenotypes have been proposed and the first set of classification criteria have been recently created. Our objectives were to assess the phenotype distribution and the performance of the classification criteria in Spanish patients as genetic and geographical differences may exist.

Methods. We performed a cross-sectional multicentre study (Registro Español de Enfermedad Relacionada con la IgG4, REERIGG4) with nine participating centres from Spain. Patients were recruited from November 2013 to December 2018. The 2019 American College of Rheumatology/European League Against Rheumatism classification criteria (AECC) were used.

Results. We included 105 patients; 88% had Caucasian ethnicity. On diagnosis, 86% met the international pathology consensus while 92% met the Japanese comprehensive criteria. The phenotype distribution was head and neck 25%, Mikulicz and systemic (MS) 20%, pancreato-hepato-biliary (PHB) 13%, retroperitoneal and aorta (RA) 26%. Sixteen per cent had an undefined phenotype. Seventy-seven per cent of the cases met the AECC. From the 24 patients not meeting the AECC, 33% met exclusion criteria, and 67% did not get a score $\geq$20 points. Incomplete pathology reports were associated to failure to meet the AECC.

Conclusions. The PHB phenotype was rare among Spanish IgG4-RD patients. The MS phenotype was less frequent and the RA phenotype was more prevalent than in other, Asian patient series. An undefined phenotype should be considered as some patients do not fall into any of the categories. Three quarters of the cases met the 2019 AECC. Incomplete pathology reports were the leading causes of failure to meet the criteria.

Key words: IgG4-related disease, classification criteria, phenotypes

CLINICAL SCIENCE

Introduction

Since IgG4-related disease (IgG4-RD) was described as a medical condition in 2001 [1], knowledge of it has rapidly increased. The most extended approach to its pathophysiological mechanism was established in 2017 [2]. The recognition of certain antigens would prompt a $CD4^+$ cytotoxic lymphocyte oligoclonal expansion [3]; an inflammatory cascade along with B cells [4] and follicular helper T cells [5] would be unleashed, leading finally to fibrosis and organ dysfunction. In the past few years, two potential antigens, galectin 3 [6] and laminin 511 [7], have been identified. From the clinical standpoint, two sets of diagnostic criteria were established in 2012, the Japanese comprehensive criteria [8] (JCC, based on

[1]Internal Medicine Department, Hospital Universitari Vall d'Hebron, Barcelona, Spain, [2]Rheumatology division and General Internal Medicine department–Windsor campus, University of Western Ontario, London, ON, Canada, [3]Internal Medicine Department, Hospital General Universitario Gregorio Marañón, Madrid, Spain, [4]Muscle Disease Unit, National Institute of Arthritis and Musculoskeletal and Skin Diseases, Bethesda, MD, USA, [5]Internal Medicine Department, Hospital Clínico Universitario de Santiago, Santiago de Compostela, [6]Internal Medicine Department, Consorci Sanitari Parc Taulí, Sabadell, [7]Internal Medicine Department, Hospital 12 de octubre, Madrid, [8]Internal Medicine Department, Hospital de Cabueñes, Gijón, [9]Internal Medicine Department, Hospital Miguel Servet, Zaragoza, [10]Internal Medicine Department, Hospital Joan XXIII, Tarragona and [11]Internal Medicine Department, Hospital Clínic i Provincial, Barcelona, Spain

Submitted 9 January 2020; accepted 15 March 2020

*Correspondence to: Andreu Fernández-Codina, Internal Medicine Department, Systemic Autoimmune Diseases Unit, Hospital Universitari Vall d'Hebron. Passeig Vall d'Hebron 119-129, PC 08035 Barcelona, Spain. E-mail: doccodina@gmail.com

Rheumatology key messages

- The pancreato-hepato-biliary phenotype was rare in a Spanish IgG4-related disease (IgG4-RD) series.
- The retroperitoneal and aorta phenotype was more common in Spanish than Asian IgG4-RD series patients.
- IgG4-RD pathology exams are not mandatory but have a significant weight in the ACR/EULAR classification criteria.

organ involvement, serum IgG4 and pathology findings) and the international pathology consensus [9] (IPC, based on pathology findings and clinical correlation). The creation of an international multiethnic cohort has allowed differentiating several phenotypes of IgG4-RD using the latent class analysis method [10]. The possibility of establishing a distinct prognosis and treatment for each phenotype is yet to be determined. IgG4-RD had a predilection for head and neck involvement in Asian patients, and differences could be due to genetic or environmental factors. Finally, the American College of Rheumatology/European League Against Rheumatism classification criteria (AECC) have just been released [11]. This is the first set of classification criteria, aiming to select patients to be enrolled in future research. They are based on clinical, radiological, laboratory and pathology findings. In the multiethnic international validation cohort, sensitivity was 85% and specificity 99.2%. The above-mentioned geographical differences could influence the performance of these classification criteria. The aim of our study is to describe the different IgG4-RD phenotypes in a mostly Caucasian group of Spanish patients and to explore the performance of the new AECC among them.

Methods

Patients

The Spanish registry of IgG4-RD (Registro Español de Enfermedad Relacionada con la IgG4, REERIGG4) was created in 2013 by the Systemic Autoimmune Diseases Group (GEAS) of the Spanish Internal Medicine Society. It was approved by the Vall d'Hebron University Hospital ethics review board. The present study included data gathered from November 2013 to December 2018. The database was an encrypted and de-identified Microsoft Access file.

Data

The dataset included the age at diagnosis of IgG4-RD, gender and ethnicity. All patients had to meet at least one of the two diagnostic criteria sets available (JCC or IPC). The different categories of the strength of the diagnosis were also recorded (definite, probable, possible for JCC; highly suggestive, probable, insufficient for IPC). All biopsies were recorded. Biopsy reports were categorized into sufficient (IgG4-RD diagnosis and report with full assessment for pathology parameters including storiform fibrosis, lymphoplasmacytic infiltrates and obliterative phlebitis, plasma cell counts for IgG4 and IgG/CD138 and IgG4/IgG ratio) or insufficient (IgG4-RD diagnosis but in the absence of any or all of the pathology parameters).

Two expert physicians (A.F.C. and F.M.V.) reviewed all cases and assigned a phenotype for each one. The possible phenotypes were pancreato-hepato-biliary (PHB), retroperitoneum and aorta (RA), head and neck limited (HN), and Mikulicz and systemic disease (MS), as described by Wallace *et al.* [10]. One extra subset (not defined, ND) was added for the cases that did not fit into one of the original four phenotypes.

The 2019 AECC are based on an entry criterion, multiple exclusion criteria and a score of points obtained from different domains (pathology, immunostaining, serum IgG4, lacrimal or salivary gland involvement, thorax involvement, pancreas and biliary tree involvement, kidney involvement, and retroperitoneal involvement). In order to diagnose a single patient with IgG4-RD according to the classification criteria, individuals needed to meet the entry criterion, to have no exclusion criteria and to have an inclusion criteria score $\geq$20 points [11]. All the criteria were introduced as individual variables in the database, and scores for each inclusion domain and total scores were documented. The 2019 AECC were based on two validation cohorts, and 39 patients were included in the first validation cohort.

Statistical analysis

Dichotomous variables were expressed as percentages and absolute frequencies, and continuous features were reported as mean (S.D.). χ^2 was used for pairwise comparisons of categorical variables between groups. Student's *t* test was used to compare continuous variables among groups. Cohen's kappa was used to measure the reliability among different sets of criteria ($>$0.75 was considered excellent, 0.4–0.75 good and $<$0.4 poor reliability). Statistical analyses were performed using SPSS Statistics (IBM Corp., Armonk, NY, USA). A two-sided *P*-value of 0.05 or less was considered statistically significant with no correction for multiple comparisons.

Results

One hundred and five IgG4-RD patients were recruited from nine Spanish hospitals. The epidemiological features and the percentage meeting each set of diagnostic criteria

for the whole group of patients are displayed in Table 1. Most (88%) of them were Caucasian. The most frequently involved organs were retroperitoneum, lymph nodes, orbits, salivary glands and pancreas. Half of the individuals had a systemic disease involving >1 organ and 42% of them had elevated serum IgG4. About 90% of the patients met IPC and/or JCC diagnostic criteria (84% met both).

Overall agreement between the two experts classifying patients by phenotype was excellent ($\kappa = 0.755$). Agreement by phenotype was excellent for HN ($\kappa = 0.898$), ND ($\kappa = 0.773$) and PHB ($\kappa = 0.922$), and good for MS ($\kappa = 0.496$). The distribution by phenotype was the following: HN 26 (25%), MS 21 (20%), ND 17 (16%), PHB 14 (13%) and RA 27 (26%) patients. The

Table 1 Demographic characteristics, organ involvement and clinical phenotypes in Spanish patients with IgG4-related disease

	HN ($n = 26$)	MS ($n = 21$)	ND ($n = 17$)	PHB ($n = 14$)	RA ($n = 27$)	Total ($n = 105$)
Age at diagnosis, mean (S.D.), years	50.1 (13.5)*	55.1 (12.6)	53.7 (15.2)	61.0 (14.9)	58.7 (12.3)	55.4 (13.8)
Female, % (*n*)	58 (15)**	33 (7)	29 (5)	29 (4)	15 (4)*	33 (35)
White, % (*n*)	88 (23)	67 (14)*	88 (15)	86 (12)	89 (24)	84 (88)
North African/Middle East, % (*n*)	4 (1)	14 (3)	0 (0)	7 (1)	0 (0)	5 (5)
Hispanic, % (*n*)	8 (2)	19 (4)	12 (2)	7 (1)	11 (3)	11 (12)
Biopsy, % (*n*)	100 (26)	95 (20)	100 (17)	64 (9)***	93 (25)	92 (97)
Pathology standard, % (*n*)	81 (21)	60 (12)	59 (10)	78 (7)	80 (20)	72 (70)
Elevated serum IgG4, % (*n*)	19 (5)**	62 (13)*	41 (7)	86 (12)***	26 (7)	42 (44)
IPC, % (*n*)	100 (26)*	90 (19)	76 (13)	50 (7)***	93 (25)	86 (90)
Highly suggestive, % (*n*)	65 (17)	62 (13)	47 (8)	43 (6)	63 (17)	58 (61)
Probable, % (*n*)	35 (9)	29 (6)	29 (5)	7 (1)	30 (8)	28 (29)
JCC, % (*n*)	96 (25)	95 (20)	88 (15)	100 (14)	93 (25)	94 (99)
Definite, % (*n*)	23 (6)	52 (11)*	18 (3)	36 (5)	33 (9)	32 (34)
Probable, % (*n*)	73 (19)**	29 (6)	41 (7)	21 (3)	48 (13)	46 (48)
Possible, % (*n*)	0 (0)**	14 (3)	29 (5)	43 (6)*	11 (3)	16 (17)
Systemic, % (*n*)	35 (9)	100 (21)***	53 (9)	43 (6)	30 (8)*	50 (53)
Pancreas, % (*n*)	0 (0)*	19 (4)	6 (1)	71 (10)***	0 (0)*	14 (15)
Lacrimal glands, % (*n*)	19 (5)*	14 (3)	0 (0)	0 (0)	0 (0)	8 (8)
Orbit, % (*n*)	50 (13)***	24 (5)	0 (0)*	0 (0)	0 (0)**	17 (18)
Extraocular muscles, % (*n*)	15 (4)	14 (3)	0 (0)	0 (0)	0 (0)	7 (7)
Salivary glands, % (*n*)	23 (6)	43 (9)***	0 (0)	7 (1)	0 (0)*	15 (16)
Pachymeninges, % (*n*)	0 (0)	24 (5)***	0 (0)	0 (0)	0 (0)	5 (5)
Hypophysis, % (*n*)	4 (1)	0 (0)	0 (0)	0 (0)	0 (0)	1 (1)
Thyroid, % (*n*)	0 (0)	0 (0)	18 (3)**	0 (0)	0 (0)	3 (3)
Aorta, % (*n*)	0 (0)*	24 (5)	0 (0)	0 (0)	30 (8)**	12 (13)
Arteries, % (*n*)	0 (0)	10 (2)	6 (1)	0 (0)	4 (1)	4 (4)
Mediastinum, % (*n*)	0 (0)	19 (4)**	6 (1)	0 (0)	0 (0)	5 (5)
Retroperitoneum, % (*n*)	0 (0)***	52 (11)	29 (5)	0 (0)**	78 (21)***	35 (37)
Mesenterium, % (*n*)	0 (0)	10 (2)	24 (4)*	0 (0)	4 (1)	7 (7)
Skin, % (*n*)	4 (1)	5 (1)	0 (0)	0 (0)	0 (0)	2 (2)
Lymph nodes, % (*n*)	8 (2)	52 (11)***	24 (4)	14 (2)	7 (2)	20 (21)
Biliary ducts, % (*n*)	0 (0)	5 (1)	0 (0)	29 (4)***	0 (0)	5 (5)
Gallbladder, % (*n*)	0 (0)	0 (0)	0 (0)	21 (3)**	0 (0)	3 (3)
Liver, % (*n*)	0 (0)	5 (1)	0 (0)	14 (2)*	0 (0)	3 (3)
Lung, % (*n*)	0 (0)	33 (7)***	18 (3)	0 (0)	0 (0)	10 (10)
Pleura, % (*n*)	0 (0)	14 (3)	6 (1)	0 (0)	4 (1)	5 (5)
Pericardium, % (*n*)	0 (0)	10 (2)	6 (1)	7 (1)	0 (0)	4 (4)
Kidney, % (*n*)	0 (0)	29 (6)***	0 (0)	0 (0)	4 (1)	7 (7)
Breast, % (*n*)	0 (0)	0 (0)	6 (1)	0 (0)	0 (0)	1 (1)
Prostate, % (*n*)	0 (0)	5 (1)	0 (0)	0 (0)	0 (0)	1 (1)
Maxillary sinus, % (*n*)	23 (6)*	24 (5)*	0 (0)	0 (0)	0 (0)	10 (11)
Other organs, % (*n*)	15 (4)	24 (5)*	6 (1)	0 (0)	0 (0)	10 (10)

Dichotomous variables were expressed as percentage (count) and continuous variables as mean (S.D.). Bivariate comparisons of continuous variables were made using Student's *t*-test while bivariate comparisons of dichotomous variables were made either using χ^2 test or Fisher's exact test, as appropriate. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$. HN: head and neck; IgG4: immunoglobulin G 4; IPC: international pathology criteria; JCC: Japanese comprehensive criteria; MS: Mikulicz and systemic; ND: not defined; PHB: pancreato-hepato-biliary; RA: retroperitoneum and aorta.

organ involvement in each subset matched each phenotype's definition.

Eighty-one (77%) met the 2019 AECC, with a mean score of 32 points (S.D. 9.7). The 24 patients not meeting these criteria (Table 2): had statistically significantly fewer biopsies, had fewer complete standard pathology reports, met the IPC diagnostic criteria with a lower frequency, and were more frequently classified as possible IgG4-RD according to the JCC criteria. All patients met the entry criterion (typical organs involved or suggestive tissue inflammation). Eight patients (33%) had at least one exclusion criterion: fever (1), steroid resistance (2), eosinophilia (1), anti-neutrophil cytoplasmic antibodies (2), anti-double stranded deoxyribonucleic acid antibodies (2), other antibodies (2), and prominent neutrophilic inflammation (1). Finally, 16 (67%), did not get a score $\geq$20 points (mean 17.2, S.D. 10). Even though the agreement between the three sets of criteria was good (all $>$73.3%), the kappa coefficients showed poor reliability between the different sets of criteria (all $<$0.4).

Discussion

The new 2019 AECC will allow selection of homogeneous populations of IgG4-RD patients for clinical trials thanks to its very high specificity. Herein, we have reported that in a non-selected Spanish IgG4-RD patient cohort (predominantly Caucasian), we found a low number of patients with the PHB (13%) phenotype and a group that did not fit in any of the prespecified phenotypes (16%). Seventy-seven per cent of the individuals met the AECC. Most of the excluded patients did not achieve the score threshold. The absence of biopsies and incomplete pathology reports accounted for a significant number of exclusions.

The existence of differences between Asian and non-Asian populations in IgG4-RD had been suggested by some observational studies. These findings classically included higher serum IgG4 levels and some concerns for increased pancreatic involvement in Asian patients [12]. After other cohorts from America, Asia and Europe were reported, the most commonly involved organs were found to be lymph nodes, submandibular and lacrimal glands, pancreas, and retroperitoneum [13–15]. The percentage of patients with elevated serum IgG4 ranged from 100% to 51%. In the longest international and multiethnic cohort available, Wallace *et al.* [10] found that Asian patients were significantly older, had higher serum IgG4 levels, and had more head and neck disease. This study included 493 patients from a derivation cohort from the AECC taskforce. Spanish patients matched in terms of age and sex this international group's results for non-Asian patients (69% Caucasian *vs* 88% in ours). In the present study a lower proportion of patients with elevated levels of serum IgG4 (69% *vs* 44%) was found. In terms of organ involvement, our Spanish cohort had more cases with rare organs involved, like hypophysis, thyroid, mesenterium or breast. Most likely, this is due to the usual clinical practice nature of REERIGG4 *vs* Wallace's series with prototypic patients in the context of the classification criteria effort. We also found a higher proportion of individuals with retroperitoneal involvement (35% *vs* 16%), despite a similar rate of aorta involvement (12% *vs* 10%). We documented less pancreatic involvement (15%), as compared with 42% in the AECC taskforce study. Moreover, all the different phenotypes had a similar overall distribution (regardless of ethnicity), except for the PHB subset. Again, only 13% of the cases accounted for the PHB phenotype. The rest of the phenotypes rarely had pancreatic, liver or bile duct involvement except for the MS phenotype (19% pancreas, 5% liver, 5% biliary ducts). Since typically multiple organs would be involved at the same time in MS, including the ones encompassed in the PHB subset, there could be some overlap between the two of them. Also, in the case of rarer manifestations, the existence of an ND phenotype makes sense and, in our dataset, represents a minority of the cases. Taking into account the phenotype sub-analysis by ethnic groups, Wallace's series showed HN 37%, MS 26%, PHB 26% and RA 14% for Asians; and HN 13%, MS 21%, PHB 32% and RA 30% for non-Asians. In our Spanish cohort (without ethnically Asian patients) the HN phenotype percentage (25%) was higher than in the non-Asian subgroup but lower than in the Asian one. The MS subset was lower (20%) than the one in Asian patients. For PHB our series showed a lower prevalence (13%) compared with both subgroups; remarkably more non-Asian patients had the PHB phenotype. Finally, the RA phenotype was more frequent in our study and in the non-Asian group *vs* the Asian group.

In terms of phenotypes, the agreement between both experts clinically assessing the phenotypes was excellent. The agreement was lower for the MS phenotype, as on some occasions, in the event of multiple organ involvement, the experts favoured either the main manifestation (i.e. pancreas and then assigning the PHB phenotype) or the summary of all the ongoing involvement (i.e. pancreas, elevated serum IgG4, retroperitoneal fibrosis, lymph nodes, and then choosing the MS phenotype). These conflicting opinions reflect regular clinical practice controversies. The findings on phenotype distribution support the presence of ethnic and geographical differences. The differences might be explained by both genetic and environmental causes yet to be determined. Ethnically, in Asian patients, the HN phenotype seems more prevalent and the RA phenotype is rarer, while serum IgG4 levels would be increased in most of the patients. The differences between non-Asian groups as different PHB prevalence could be determined by geographical causes such as exposure to different antigens that may cause different manifestations of the same disease.

Since there is no previous set of IgG4-RD classification criteria, it is difficult to choose a comparator to assess the AECC. In the original second validation cohort [11], sensitivity and specificity were remarkably high (82% and 97.8%, respectively). In our study, 81 of the

Table 2 Characteristics of patients according to the fulfilment of the 2019 IgG4-related disease ACR/EULAR classification criteria

	Meets 2019 IgG4-RD ACR/EULAR classification criteria			
	No ($n = 24$)	Yes ($n = 81$)	*P*-value	Total ($n = 105$)
Age at diagnosis, mean (S.D.), years	55.5 (15.1)	55.3 (13.5)	1.0	55.4 (13.8)
Female, % (*n*)	21 (5)	37 (30)	0.1	33 (35)
White, % (*n*)	100 (24)	79 (64)	0.01	84 (88)
North African/Middle East, % (*n*)	0 (0)	6 (5)	0.6	5 (5)
Hispanic, % (*n*)	0 (0)	15 (12)	0.06	11 (12)
Biopsy, % (*n*)	79 (19)	96 (78)	0.01	92 (97)
Pathology standard, % (*n*)	42 (8)	79 (62)	0.001	72 (70)
Elevated serum IgG4, % (*n*)	58 (14)	37 (30)	0.06	42 (44)
IPC, % (*n*)	62 (15)	93 (75)	<0.001	86 (90)
Highly suggestive, % (*n*)	33 (8)	65 (53)	0.005	58 (61)
Probable, % (*n*)	29 (7)	27 (22)	0.8	28 (29)
JCC, % (*n*)	96 (23)	94 (76)	1.0	94 (99)
Definite, % (*n*)	25 (6)	35 (28)	0.4	32 (34)
Probable, % (*n*)	33 (8)	49 (40)	0.2	46 (48)
Possible, % (*n*)	38 (9)	10 (8)	0.003	16 (17)
Systemic, % (*n*)	46 (11)	52 (42)	0.6	50 (53)
Pancreas, % (*n*)	12 (3)	15 (12)	1.0	14 (15)
Lacrimal glands, % (*n*)	4 (1)	9 (7)	0.7	8 (8)
Orbit, % (*n*)	12 (3)	19 (15)	0.8	17 (18)
Extraocular muscles, % (*n*)	4 (1)	7 (6)	1.0	7 (7)
Salivary glands, % (*n*)	17 (4)	15 (12)	0.8	15 (16)
Pachymeninges, % (*n*)	4 (1)	5 (4)	1.0	5 (5)
Hypophysis, % (*n*)	0 (0)	1 (1)	1.0	1 (1)
Thyroid, % (*n*)	0 (0)	4 (3)	1.0	3 (3)
Aorta, % (*n*)	12 (3)	12 (10)	1.0	12 (13)
Arteries, % (*n*)	0 (0)	5 (4)	0.6	4 (4)
Mediastinum, % (*n*)	0 (0)	6 (5)	0.6	5 (5)
Retroperitoneum, % (*n*)	25 (6)	38 (31)	0.2	35 (37)
Mesenterium, % (*n*)	8 (2)	6 (5)	0.7	7 (7)
Skin, % (*n*)	0 (0)	2 (2)	1.0	2 (2)
Lymph nodes, % (*n*)	21 (5)	20 (16)	1.0	20 (21)
Biliary ducts, % (*n*)	12 (3)	2 (2)	0.08	5 (5)
Gallbladder, % (*n*)	0 (0)	4 (3)	1.0	3 (3)
Liver, % (*n*)	0 (0)	4 (3)	1.0	3 (3)
Lung, % (*n*)	17 (4)	7 (6)	0.2	10 (10)
Pleura, % (*n*)	12 (3)	2 (2)	0.08	5 (5)
Pericardium, % (*n*)	4 (1)	4 (3)	1.0	4 (4)
Kidney, % (*n*)	0 (0)	10 (7)	0.3	7 (7)
Breast, % (*n*)	0 (0)	1 (1)	1.0	1 (1)
Prostate, % (*n*)	0 (0)	1 (1)	1.0	1 (1)
Maxillary sinus, % (*n*)	17 (4)	9 (7)	0.3	10 (11)
Other organs, % (*n*)	8 (2)	10 (8)	0.7	10 (10)

Dichotomous variables are expressed as percentage (count) and continuous variables as mean (S.D.). Bivariate comparisons of continuous variables were made using Student's *t*-test while bivariate comparisons of dichotomous variables were made either using χ^2 test or Fisher's exact test, as appropriate. IgG4: immunoglobulin G 4; IgG4-RD ACR/EULAR: IgG4-related disease American College of Rheumatology/European League Against Rheumatism; IPC: international pathology criteria; JCC: Japanese comprehensive criteria.

IgG4-RD cases (77%) met the classification criteria. All patients met the inclusion criteria. It is important to remark that patients can have the characteristic clinical or radiological involvement from typical organs, but that also organs with biopsy proven lymphoplasmacytic infiltrate of uncertain origin can be included. This alternative proven biopsy allowed us to include 10 patients that we had initially discarded since they had rarer manifestations not specified in the list provided making a restrictive interpretation of the 2019 AECC. The exclusion criteria, which encompass multiple mimicker conditions, excluded just a few patients, mostly due to serologic

causes. Some of them had positive antibodies like ANCA, without evidence of active vasculitis. Case series have suggested the possibility of an overlap syndrome between ANCA vasculitis and IgG4-RD, but it is still controversial [16]. Most likely clinical manifestations would tilt the balance towards one disease or the other. This could also happen with other autoimmune conditions. Nevertheless, classification criteria should select the most representative and homogeneous types of patients to improve the homogeneity of clinical trials, and this is why cases with any potential bias should not go through. Exclusion criteria also include important key points such as excluding patients with malignancies. In previous studies, IgG4-RD had been related to synchronic diagnosis with cancer and early development of malignancies in the 2 years after diagnosis [17, 18]. These were Asian patients fulfilling JCC, mostly without biopsies. By prompting broader studies in the case of unexplained or suspicious masses, and directly excluding patients with biopsy-proven malignancy even in coexistence with IgG4-RD features, cancer development rates in IgG4RD might drop. Future epidemiological studies using the AECC might be able to clarify the real association between IgG4-RD and cancer. Classification criteria can be achieved without having a pathology sample or serum IgG4 determinations. Patients not achieving the 20-point threshold (and therefore not meeting the classification criteria) were found to lack a biopsy sample more often (79% *vs* 96%). Ninety per cent of the patients at REERIGG4 had biopsies performed for diagnosis. We also found that having a biopsy done was a significant item in terms of meeting the classification criteria. Beyond that, among those who already had a biopsy, having a complete pathology report including all the standard features and counts [9] made the difference. Patients with incomplete information had lower scores. This emphasizes the importance of having accurate pathology reports.

Lastly, patients who met the AECC met more often the IPC diagnostic criteria than those who did not. Moreover, those in the highly suggestive category also met more frequently the classification criteria. This is most likely due to the significant weight of pathology in the scoring system since the IPC are based on pathology and patients with more items (high $IgG4^+$/IgG4 ratios, all classic histological features) will certainly have higher scores. The JCC were equally met by cases meeting and not meeting the AECC. Remarkably, more patients who did not make it for the classification criteria fell in the lower certainty of diagnosis category (possible), which does not warrant a biopsy sample. In the absence of an older set of IgG4-RD classification criteria as a comparator, we exploratorily compared the new AECC agreement with the IPC and JCC diagnostic criteria without getting any correlation.

The limitations of this study include the absence of a central laboratory/pathology department to process all the samples in a single place, which is inherent to a multicentre rare disease registry. In addition, all the participant clinics are specialized in General Internal Medicine and systemic autoimmune diseases. Thus, some manifestations might be under-represented as pancreatic involvement. However, as a systemic disease, patients tend to have multiple organs involved at the same time and Internal Medicine is involved anyway. Some cases were previously used for the creation of a derivation case collection for the 2019 AECC. Finally, the phenotype assignment by two experts might include biases as compared with the use of a computer algorithm but seems to us closer to clinical practice.

In conclusion, we found that the PHB phenotype was rare in an IgG4-RD Spanish multicentre cohort, mostly formed by Caucasian patients. The MS phenotype was less frequent and the RA phenotype was more frequent than in other, Asian patient series. Some cases with less frequent manifestations might warrant the addition of a ND phenotype. Seventy-seven per cent of the patients met the 2019 AECC. Lack of biopsies and incomplete pathology reports were the leading causes of failure to meet the criteria. Pathology was still determinant although not mandatory thanks to the AECC design. These criteria will help in homogenizing populations of IgG4-RD patients for future clinical trials and epidemiological studies.

Acknowledgements

All authors are the members of the registry 'Registro Español de Enfermedad Relacionada con la IgG4' (REERIGG4) and would like to acknowledge the Autoimmune Systemic Diseases Group (GEAS) within the Spanish Internal Medicine Society (SEMI) for its support.

Funding: No specific funding was received from any funding bodies in the public, commercial or not-for-profit sectors to carry out the work described in this manuscript.

Disclosure statement: The authors have declared no conflicts of interest.

References

1 Hamano H, Kawa S, Horiuchi A *et al* High serum IgG4 concentrations in patients with sclerosing pancreatitis. N Engl J Med 2001;344:732–8.

2 Mattoo H, Stone JH, Pillai S. Clonally expanded cytotoxic $CD4^+$ T cells and the pathogenesis of IgG4-related disease. Autoimmunity 2017;50:19–24.

3 Della-Torre E, Bozzalla-Cassione E, Sciorati C *et al* A $CD8\alpha-$ subset of CD4+SLAMF7+ cytotoxic T cells is expanded in patients with IgG4-related disease and decreases following glucocorticoid treatment. Arthritis Rheumatol 2018;70:1133–43.

4 Lanzillotta M, Della-Torre E, Milani R *et al* Increase of circulating memory B cells after glucocorticoid-induced remission identifies patients at risk of IgG4-related disease relapse. Arthritis Res Ther 2018;20:222.

5 Akiyama M, Yasuoka H, Yamaoka K *et al* Enhanced IgG4 production by follicular helper 2 T cells and the involvement of follicular helper 1 T cells in the pathogenesis of IgG4-related disease. Arthritis Res Ther 2016;18:167.

6 Perugino CA, AlSalem SB, Mattoo H *et al* Identification of galectin-3 as an autoantigen in patients with IgG4-related disease. J Allergy Clin Immunol 2019;143:736–45.e6.

7 Shiokawa M, Kodama Y, Sekiguchi K *et al* Laminin 511 is a target antigen in autoimmune pancreatitis. Sci Transl Med 2018;10:eaaq0997.

8 Umehara H, Okazaki K, Masaki Y *et al* Comprehensive diagnostic criteria for IgG4-related disease (IgG4-RD), 2011. Mod Rheumatol 2012;22:21–30.

9 Deshpande V, Zen Y, Chan JK *et al* Consensus statement on the pathology of IgG4-related disease. Mod Pathol 2012;25:1181–92.

10 Wallace ZS, Zhang Y, Perugino CA *et al* Clinical phenotypes of IgG4-related disease: an analysis of two international cross-sectional cohorts. Ann Rheum Dis 2019;78:406–12.

11 Wallace ZS, Naden RP, Chari S *et al* The 2019 American College of Rheumatology/European League Against Rheumatism Classification criteria for IgG4-related disease. Arthritis Rheumatol 2020:72:7–19.

12 Inoue D, Yoshida K, Yoneda N *et al* IgG4-related disease: dataset of 235 consecutive patients. Medicine (Baltimore) 2015;94:e680.

13 Martínez-Valle F, Fernández-Codina A, Pinal-Fernández I, Orozco-Gálvez O, Vilardell-Tarrés M. IgG4-related disease: evidence from six recent cohorts. Autoimmun Rev 2017;16:168–72.

14 Campochiaro C, Ramirez G, Bozzolo E *et al* IgG4-related disease in Italy: clinical features and outcomes of a large cohort of patients. Scand J Rheumatol 2016;45:135–45.

15 Ebbo M, Daniel L, Pavic M *et al* IgG4-related systemic disease: features and treatment response in a French cohort: results of a multicenter registry. Medicine (Baltimore) 2012;91:49–56.

16 Danlos F-X, Rossi GM, Blockmans D *et al* Antineutrophil cytoplasmic antibody-associated vasculitides and IgG4-related disease: a new overlap syndrome. Autoimmun Rev 2017;16:1036–43.

17 Asano J, Watanabe T, Oguchi T *et al* Association between immunoglobulin G4-related disease and malignancy within 12 years after diagnosis: an analysis after longterm followup. J Rheumatol 2015;42:2135–42.

18 Yamamoto M, Takahashi H, Tabeya T *et al* Risk of malignancies in IgG4-related disease. Mod Rheumatol 2012;22:414–8.