

OPEN

# Diagnostic Value of Serum IgG4 for IgG4-Related Disease

## A PRISMA-compliant Systematic Review and Meta-analysis

Mingju Hao, MD, Min Liu, MM, Gaowei Fan, MD, Xin Yang, MD, and Jinming Li, PhD

**Abstract:** Many studies about serum IgG4 for the diagnosis of IgG4-related disease (IgG4-RD) have been reported. However, these studies had relatively small sample sizes and the diagnostic accuracy values varied much between them.

The aim of this study was to perform a meta-analysis to evaluate the diagnostic value of serum IgG4 for IgG4-RD.

We conducted a search of relevant articles using MEDLINE, EMBASE, Web of Science, SCOPUS, and Cochrane Library databases published before December 2015.

Studies those assessed the diagnostic accuracy of serum IgG4 for IgG4-RD and those provided the cut-off value for serum IgG4 were included.

Data were synthesized using the random-effect model. Statistical analysis was performed using STATA with the MIDAS module and Meta-DiSc 1.4 software.

A total of 9 case-control studies were analyzed, which included 1235 patients with IgG4-RD and 5696 overall controls. The pooled estimate, for a cut-off value ranged from 135 to 144 mg/dL, produced a sensitivity of 87.2% (95% CI, 85.2–89.0%) and a specificity of 82.6% (95% CI, 81.6–83.6%). The positive likelihood ratio (PLR), negative likelihood ratio (NLR), and diagnostic odds ratio (DOR) were 6.48 (95% CI, 3.98–10.57), 0.14 (95% CI, 0.09–0.21), and 45.15 (95% CI, 23.41–87.06), respectively. The area under the curve (AUC) of the summary receiver operating characteristic curve (SROC) was 0.94 (0.92–0.96). When a cut-off value of 2-fold the upper limit of normal was used (ranged from 270 to 280 mg/dL), the pooled sensitivity was 63% (95% CI, 60.0–66.0%), and the specificity was 94.8% (95% CI, 94.1–95.4%). The PLR, NLR, and DOR were 13.3 (95% CI, 7.39–24.0), 0.41 (95% CI, 0.29–0.58) and 33.42 (95% CI, 13.88–80.43), respectively. The AUC of the SROC was 0.92 (0.90–0.94).

Only a relatively small number of studies were included, and significant heterogeneity was observed in this meta-analysis.

Serum IgG4 is a modestly effective marker to diagnose IgG4-RD. Doubling the cut-off value for IgG4 could not improve the overall test characteristics. A high specificity inevitably accompanies with a significant sacrifice in sensitivity.

Editor: Seiho Nagafuchi.

Received: March 4, 2016; revised: April 27, 2016; accepted: May 4, 2016. From the National Center for Clinical Laboratories (MH, GF, XY, JL), Beijing Hospital; Graduate School (MH, GF, XY, JL), Peking Union Medical College, Chinese Academy of Medical Sciences, Beijing; Department of Clinical Laboratory (MH), Qianfo Mountain Hospital of Shandong University; and Department of Clinical Laboratory (ML), Jinan Dermatitis Prevention and Control Hospital, Jinan, People's Republic of China.

Correspondence: Jinming Li, National Center for Clinical Laboratories, Beijing Hospital, Dongdan, Beijing, People's Republic of China (e-mail: jml@ncccl.org.cn).

The authors have no funding and conflicts of interest to disclose.

Copyright © 2016 Wolters Kluwer Health, Inc. All rights reserved.

This is an open access article distributed under the Creative Commons Attribution License 4.0, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. ISSN: 0025-7974

DOI: 10.1097/MD.0000000000003785

(*Medicine* 95(21):e3785)

**Abbreviations:** AIP = autoimmune pancreatitis, AUC = area under the curve, DOR = diagnostic odds ratio, FN = false negative, FP = false positive, IgG4-RD = IgG4-related disease, NLR = negative likelihood ratio, PLR = positive likelihood ratio, SROC = summary receiver operating characteristic curve, TN = true negative, TP = true positive.

### INTRODUCTION

IgG4-related disease is an immune-mediated condition that can affect almost any organ.<sup>1</sup> The characteristic features include a lymphoplasmacytic infiltrate composed of IgG4 positive plasma cells, storiform fibrosis, obliterative phlebitis, and mild-to-moderate eosinophilia.<sup>1</sup> Most patients with IgG4-RD have elevated serum IgG4 concentrations and abundant accumulation of IgG4-positive plasma cells in affected tissues.<sup>2</sup>

Elevated serum levels of IgG4 provide a noninvasive method to diagnose IgG4-RD. The comprehensive diagnostic criteria for IgG4-RD contained 3 major criteria: (1) clinical characteristics of diffuse/localized swelling or masses in single or multiple organs; (2) serum IgG4 concentrations  $\geq 135$  mg/dL; (3) histopathologic characteristics of marked lymphocyte and plasmacyte infiltration and fibrosis, infiltration of IgG4-positive plasma cells:  $>40\%$  of IgG4/IgG positive plasma cells and  $>10$  IgG4 positive cells/high powered field of biopsy sample.<sup>3,4</sup> Early reports associated IgG4-RD with an elevated serum IgG4 concentration demonstrated at least one-third of patients have a normal serum IgG4 concentration.<sup>5–7</sup> These findings force us to question the actual significance of IgG4 in this condition.

So far, a wide range of diagnostic accuracy values of serum IgG4 for the diagnosis of IgG4-RD have been reported, and these studies have had relatively small sample sizes.<sup>7–15</sup> Some previous meta-analysis studies focused on autoimmune pancreatitis (AIP)-the typical IgG4-RD.<sup>16,17</sup> But these AIP patients were not clearly described to be with type-1 or type-2 AIP. In order to better know the diagnosis accuracy associated with serum IgG4 concentration for the diagnosis of IgG4-RD, we reviewed the current evidence to perform this meta-analysis and systematic review.

### METHODS

#### Literature Search

We searched MEDLINE, EMBASE, Web of Science, SCOPUS, and Cochrane Library from inception to December 2015. We used the terms “IgG4-related disease” OR “IgG4-related diseases” as text words combined with “serum IgG4” OR “serum immunoglobulin g4.” We also contacted the corresponding authors by e-mail for those unpublished data of some literatures. Studies were excluded if no response was received. All analyses of this systemic review were based on previous published studies, so no ethical approval and patient consent are required.

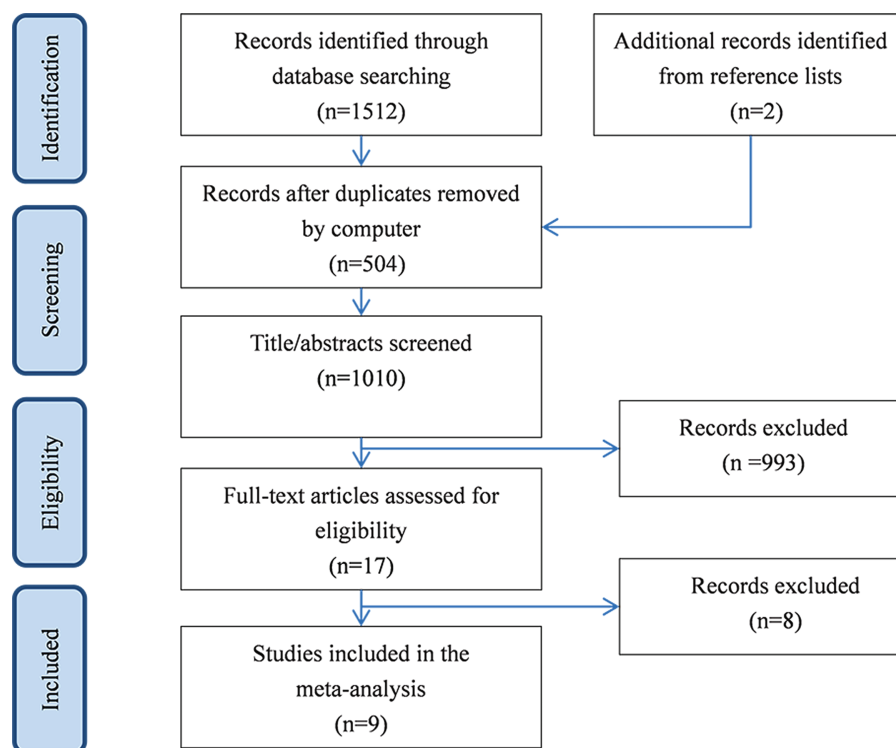


FIGURE 1. Flowchart of study selection by using the electronic database.

## Inclusion and Exclusion Criteria

Studies meeting the following criteria were eligible for inclusion: those that assessed the diagnostic accuracy of serum IgG4 for IgG4-RD; those providing detailed diagnostic criteria of IgG4-RD; sufficient data reported to construct a two-by-two table; studies providing the cut-off value for serum IgG4; no language restrictions. The exclusion criteria were: animals, cell cultures, reviews or case reports; duplicate articles; not original articles; <10 IgG4-RD patients or disease controls; insufficient data to construct a two-by-two table; studies focused on AIP which were not clearly described to be type-1 or type-2 were also excluded.

## Data Extraction

The following data were extracted from the included studies: first author; published year; geographical region; disease names and main control diseases; detection method; cut-off value; diagnostic criteria of IgG4-RD; sample size; true positive (TP); false positive (FP); false negative (FN); and true negative (TN). According to the media of sample size, eligible studies were classified as large ( $\geq$  median sample size) or small ( $<$  median sample size). The accuracy of the data was verified by a second investigator (Liu). Discrepancies between 2 reviewers were resolved by consensus after discussion. The quality of each study was assessed according to the QUADAS-2 (quality assessment of diagnostic accuracy studies 2) by 2 investigators.<sup>18</sup> Summary of QUADAS-2 plot was generated by the Review Manager Software (version 5.3, the Cochrane Collaboration).

## Statistical Analysis

The pooled sensitivity, specificity, PLR, NLR, DOR, and their confidence intervals (95% CIs) were calculated by the accuracy data (TP, FP, FN, and TN) extracted from each eligible

study. The SROC was generated, and the AUC was calculated to represent the overall performance of the detection method.

Spearman rank correlation was used to test for threshold effects, and a  $P$  value  $< 0.05$  indicated a significant threshold effect. The heterogeneity across included studies caused by the nonthreshold effect was examined by Cochran Q-statistic and  $I^2$  tests. The  $I^2$  test was used to estimate the extent of heterogeneity. A  $I^2$  value  $>30\%$  would indicate moderate heterogeneity and a value  $>50\%$  would represent substantial heterogeneity.<sup>19</sup> Possible sources of heterogeneity and their effects on the results were explored by subgroup analyses and sensitivity analyses. Prespecified subgroup analyses stratified by sample size, regions, diseases were performed to assess the influence of the above parameters. Deeks' funnel plot was performed to detect the publication bias. A  $P$  value  $<0.10$  indicated significant publication bias.

Statistical analysis was performed using the STATA software (version 12.0, Stata Corporation, TX) with the MIDAS module and Meta-DiSc 1.4 (XI, Cochrane Colloquium, Barcelona, Spain). All reported  $P$  values were 2 sided and  $P < 0.05$  was considered to be statistically significant.

## RESULTS

### Literature Search

A total of 1514 records were retrieved from initial search on electronic databases, with 321 from EMBASE, 304 from PUBMED, 640 from SCOPUS, 247 from Web of Science, and 2 from the reference lists by hand research. In total, 504 records were removed after duplicates by computer. After title and abstract screening, 993 articles were excluded. Overall, 17 full texts were examined for the eligibility of inclusion. Also, 8 articles were further excluded because they were duplicates,

TABLE 1. Characteristics and Data Extracted From the Studies Included in the Meta-Analysis

Reference	First Author	Publication Year	Region	Disease	Main Controls	Criteria
7	Carruthers et al	2015	USA	IgG4-RD	Chronic sinusitis; recurrent pneumonia; connective tissue disease; immunodeficiency; infectious, pancreatobiliary and interstitial lung disease; aortitis; leukemia; miscellaneous disorders	Consensus statement on the pathology of IgG4-RD
8	Yamamoto et al	2012	Japan	IgG4-RD	Rheumatic diseases; bronchial asthma; idiopathic pulmonary fibrosis; chronic hepatitis; liver cirrhosis; multicentric Castleman's disease; healthy individuals	HISORT criteria
9	Oseini et al	2011	USA	IgG4-associated cholangitis	Cholangiocarcinoma	HISORT criteria
10	Boonstra et al	2014	Netherlands	IgG4-Associated Cholangitis	Primary sclerosing cholangitis	HISORT criteria
11	Ohara et al	2013	Japan	IgG4-related sclerosing cholangitis	Pancreatic cancer; primary sclerosing cholangitis; cholangiocarcinoma	2012 Japanese clinical diagnostic criteria for IgG4-SC
12	Li et al	2015	China mainland	IgG4-RD	Autoimmune disease; nephropathy; cancer; pancreatitis; healthy controls	Comprehensive diagnostic criteria for IgG4-RD
13	Yu et al	2015	China Taiwan	IgG4-RD	Autoimmune diseases; malignancy; allergic disease; pancreatitis; cholangitis; miscellaneous diseases	Comprehensive diagnostic criteria for IgG4-RD
14	Su et al	2015	China mainland	IgG4-RD	Connective, digestive, hematological, and neurological disorders; carcinoma; nephropathy; Infection; gynecopathy, cardiovascular, and respiratory disorders	Comprehensive diagnostic criteria for IgG4-RD
15	Masaki et al	2012	Japan	IgG4-RD	Sjogren's syndrome, sarcoidosis, multicentric Castleman's disease, B-cell lymphoma	Comprehensive diagnostic criteria for IgG4-RD

Reference	Patient Population			Control Population			Detection Methods	Cut-off Value (mg/dL)	IgG4-RD vs Overall Controls			
	Sample Size	Male Gender	Mean Age, y	Sample Size	Male Gender	Mean Age, y			TP	FP	FN	TN
7	72	NR	NR	308	NR	NR	Nephelometry	135	65	125	7	183
8	102	NR	NR	299	NR	NR	Nephelometry	270	25	28	47	280
9	97	75	58	287	175	60	Nephelometry	144	97	28	5	271
10	73	62	64.1	310	170	51.2	Nephelometry	140	69	37	28	250
11	344	273	65.2	504	Not clear	Not clear	Nephelometry	280	51	7	22	303
12	242	154	56.5	255	161	48.6	Nephelometry	135	309	41	35	463
13	161	113	58.4	2740	1105	53.5	Nephelometry	270	235	8	109	496
14	12	NR	NR	945	NR	NR	ELISA	135	150	16	92	239
15	132	NR	NR	48	NR	NR	NR	270	121	161	40	2579
								135	12	32	0	913
								135	128	10	4	38

ELISA = enzyme linked immunosorbent assay, FN = false negative, FP = false positive, HISORt = histology, imaging, serology, other organ involvement, and response to steroid therapy criteria, IgG4-RD = Immunoglobulin G4-related disease, NR = no report, TN = true negative, TP = true positive.

review articles, or insufficient data to construct a two-by-two table. Finally, 9 case-control studies met the inclusion criteria (Figure 1).

**Characteristics of the Included Studies**

Table 1 presents the main characteristics of the included studies, as well as the diagnostic data of serum IgG4 extracted from each of the studies and analyzed. A total of 9 studies were analyzed, which included 1235 patients with IgG4-RD and 5696 healthy controls. The publication year of the included studies was 2011 or later. With regard to the geographic region of the studies, 2 were carried out in the United States, 1 in Netherland, 3 in Japan, and 3 in China. Figure 2 presents the summary plot of the QUADAS-2. As shown, the methodological quality of the eligible studies was adequate and not significantly affected by bias.

**Diagnostic Accuracy of Serum IgG4**

**Meta-Analysis for a Cut-Off Value of the Upper Limit of Normal**

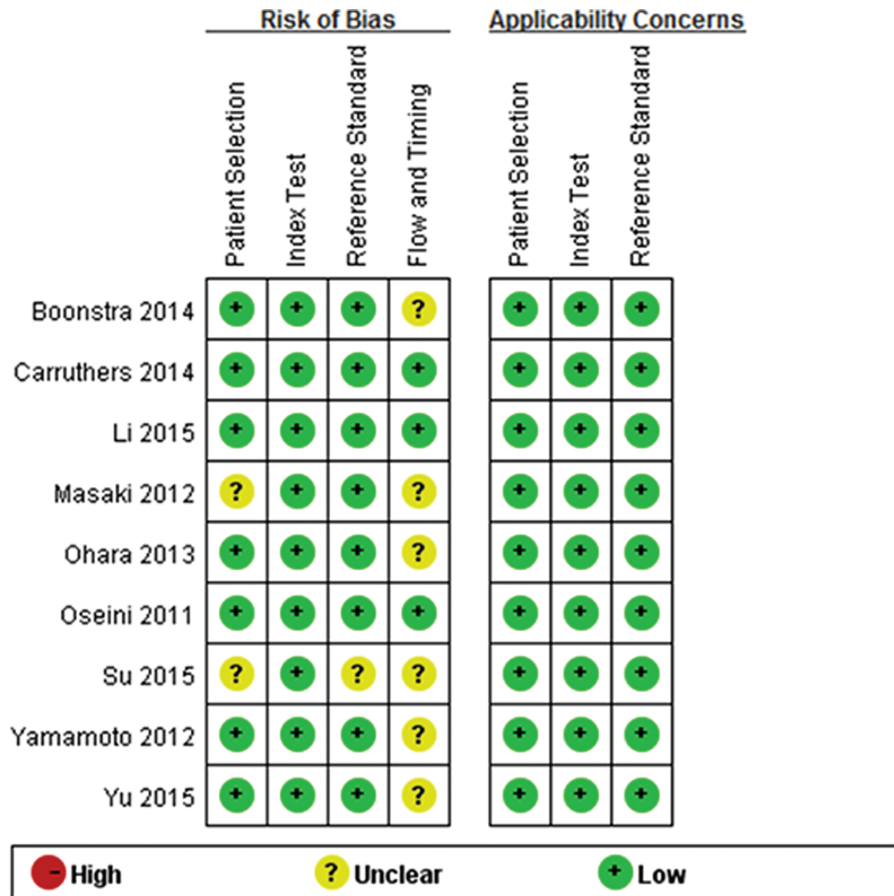
All the 9 analyzed studies provided a cut-off value of the upper limit of normal between 135 and 144 mg/dL. The reported sensitivity of the serum IgG4 for the diagnosis of IgG4-RD ranged from 71.1% to 100%, and the reported specificity ranged from 59.4% to 96.6%. The pooled estimate, by using a random-effects model, produced an average sensitivity of 87.2% (95% CI, 85.2–89.0%;  $Q = 56.2, P < 0.01; I^2 = 85.8\%$ ) and an average specificity of 82.6% (95% CI, 81.6–83.6%;  $Q = 390.57, P < 0.01; I^2 = 98\%$ ). Additionally, the average PLR and NLR were 6.48 (95% CI, 3.98–10.57;  $Q = 311.3, P < 0.01; I^2 = 97.4\%$ ) and 0.14 (95% CI, 0.09–0.21;  $Q = 50.77, P < 0.01; I^2 = 84.2\%$ ), respectively. The DOR was 45.15 (95% CI, 23.41–87.06;  $Q = 64.53, P < 0.01; I^2 = 87.6\%$ ). The AUC of the SROC was 0.94 (0.92–0.96). The sensitivity and specificity forest plots and SROC curve were shown in Figure 3A and B, and Figure 4A, respectively.

**Meta-Analysis for a Cut-Off Value of 2-Fold the Upper Limit of Normal**

In total, 6 of 9 studies in our meta-analysis provided the diagnostic values of 2-fold the upper limit of normal between 270 and 280 mg/dL. These studies included a total of 989 patients with IgG4-RD and 4404 overall controls. The reported sensitivity of the serum IgG4 for diagnosis of IgG4-RD ranged from 34.7% to 75.2%, and the reported specificity ranged from 90.9% to 98.4%. The pooled estimate, by using a random-effects model, produced an average sensitivity of 63% (95% CI, 60.0–66.0%;  $Q = 57.4, P < 0.01; I^2 = 91.3\%$ ) and an average specificity of 94.8% (95% CI, 94.1–95.4%;  $Q = 36.9, P < 0.001; I^2 = 86.5\%$ ). Additionally, the average PLR and NLR were 13.3 (95% CI, 7.39–24.0;  $Q = 48.51, P < 0.001; I^2 = 89.7\%$ ) and 0.41 (95% CI, 0.29–0.58;  $Q = 90.16, P < 0.001; I^2 = 94.5\%$ ), respectively. The DOR was 33.42 (95% CI, 13.88–80.43;  $Q = 60.51, P < 0.001; I^2 = 91.7\%$ ) and AUC of the SROC was 0.92 (0.90–0.94). The forest plots and SROC were shown in Figure 3C and D, and Figure 4B, respectively.

**Threshold Effect and Heterogeneity**

No significant threshold effect was found by visual assessment of SROC curve (Figure 4A and B). The Spearman correlation coefficient was further performed to confirm the



**FIGURE 2.** Quality assessment of included studies based on the QUADAS-2. QUADAS-2 = quality assessment of diagnostic accuracy studies 2.

effect. The Spearman correlation coefficient was  $-0.167$  ( $P = 0.668$ ) and  $-0.486$  ( $P = 0.329$ ) for the 2 cut-offs of serum IgG4, respectively, which suggested the absence of threshold effect. Considering the significant heterogeneity among studies, we performed subgroup analyses to explore the source of heterogeneity. Sample size, regions, and diseases were analyzed for each accuracy data. However, none of these above was identified as a possible source of heterogeneity. Nevertheless, the diagnostic accuracy data were consistent across different subgroups (Table 2).

**Sensitivity Analysis and Publication Bias**

Significant heterogeneity could be detected from the results of diagnostic accuracy tests. We did sensitivity analysis by omitting a single study, but the results showed the pooled results were not significantly affected by individual studies, indicating a good stability of the meta-analysis (Table 3). As shown in Figure 5A and B, Deeks’ funnel plots were almost symmetric and  $P$  values were 0.57 and 0.17 for the 2 cut-off values of serum IgG4, respectively, which suggested a low likelihood of publication bias.

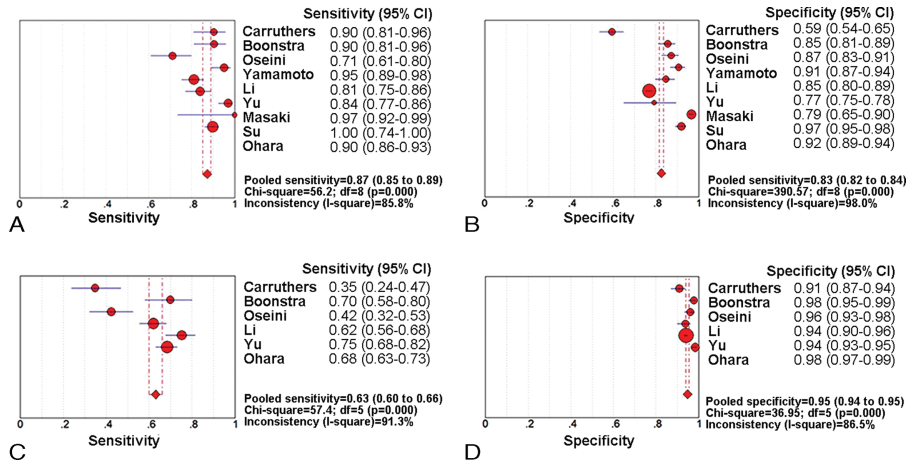
**DISCUSSION**

A serum IgG4 concentration  $>135$  mg/dL has been widely accepted as the cut-off value for the diagnosis of IgG4-RD. In this analysis, all the 9 studies provided the diagnostic values of

the upper limit of normal. By pooling the data from these studies that included a total of 6931 subjects, we calculated the pooled sensitivity for serum IgG4 in the diagnosis of IgG4-RD was 0.872 and the specificity was 0.826. The DOR is a single indicator of test performance derived from sensitivity and specificity, which is defined as the ratio of the odds of positivity in disease relative to the odds of positivity in the nondiseased.<sup>20</sup> Higher values indicate better discriminatory test performance. Our analysis showed the pooled DOR value was 45.15, which means that the serum IgG4 is a useful biomarker in diagnosing IgG4-RD. In addition, the AUC of SROC with a value of 0.94 also showed good accuracy of serum IgG4.

Considering that this cut-off level displayed lower specificity in discriminating IgG4-RD, we also performed a meta-analysis of serum IgG4 at a cut-off value of  $>2$  times the upper limit of normal. Among 9 studies included in our meta-analysis, 6 provided the diagnostic accuracy. The pooled specificity of serum IgG4 was higher (94.8%), but the sensitivity was lower (63%). The pooled DOR was 33.42 and the AUC of SROC was 92.17, which were not improved in comparison with the conventional cut-off value. The higher specificity derived from the higher cut-off value for serum IgG4 concentration, however, came at a significant reduction in sensitivity.

In addition, 2 of the enrolled studies investigated the diagnostic value of serum IgG4 with total serum IgG (IgG4/IgG) ratios for IgG4-RD. One found that the ROC curve of serum IgG4/IgG ratio was almost identical to the serum IgG4

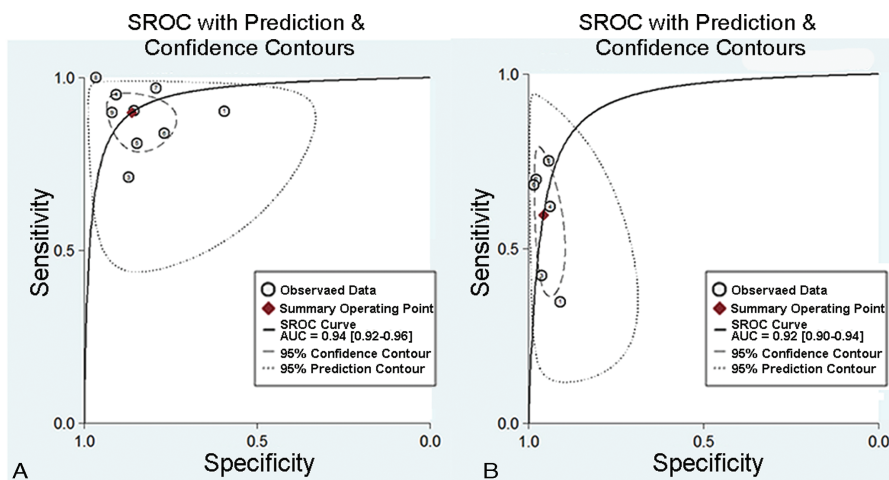


**FIGURE 3.** Forest plots of the sensitivity and specificity of serum IgG4 for IgG4-RD diagnosis. Sensitivity (A) and specificity (B) forest plots for a cut-off value of the normal upper limit. Sensitivity (C) and specificity (D) forest plots for a cut-off value of 2-fold the normal upper limit. IgG4-RD = IgG4-related disease.

concentration.<sup>15</sup> The other study by Carruthers et al<sup>7</sup> showed neither the specificity nor the positive predictive value was improved by the serum IgG4/IgG ratio for the diagnosis of IgG4-RD. These finding supported that the IgG4/IgG ratio could not improve the overall test characteristics compared with serum IgG4 alone. To better elucidate the usefulness of serum IgG4/IgG ratio, further studies are needed.

Substantial heterogeneity in the pooled accuracy data was found among studies. Subgroup analysis was performed according to sample size, regions, and diseases. But unfortunately, none of the analyzed covariates was the source of heterogeneity. We also performed sensitivity analyses and found that the pooled results were relatively stable and not affected by a single study. As we know, IgG4-RD is a systematical disorder involving numerous organs including the pancreas, salivary glands, retroperitoneum, kidneys, lymph nodes, and others. Additionally, different case controls were enrolled according to the purposes and the interested diseases. The heterogeneity could possibly be caused by difference in selection of IgG4-RD cases and controls.

Our findings were strengthened by excluding the studies concerning the diagnostic value of serum IgG4 for AIP, in which the subtypes of AIP patients were not clearly described. AIP could be classified into 2 types according to the international consensus diagnostic criteria (ICDC).<sup>21</sup> Only type 1 AIP is part of the spectrum of IgG4-RD. In a previous meta-analysis, Morselli-Labate and Pezzilli<sup>16</sup> reported the pooled sensitivity and specificity of serum IgG4 in diagnosing AIP were 82.4% and 94.6%, respectively.<sup>16</sup> Similar findings were reported in a recent study by Lian et al.<sup>17</sup> The results indicated that serum IgG4 exhibited high specificity (94%) and relatively low sensitivity (74%). However, the subtypes of these AIP patients in the included studies were not clearly described, which might explain the difference to our findings. Furthermore, although no publishing years were set during our literature searching, only publications after 2011 were included in this meta-analysis. This was consistent with that the newly emerged disease was referred to as IgG4-RD only more than a decade ago.<sup>22</sup> More detailed diagnostic criteria were proposed



**FIGURE 4.** The SROCs of serum IgG4 for IgG4-RD diagnosis. (A) SROC for a cut-off value of the normal upper limit. (B) SROC for a cut-off value of 2-fold the normal upper limit. IgG4-RD = IgG4-related disease, SROC = summary receiver operating characteristic curve.

**TABLE 2.** Subgroup Analysis of the 9 Included Studies

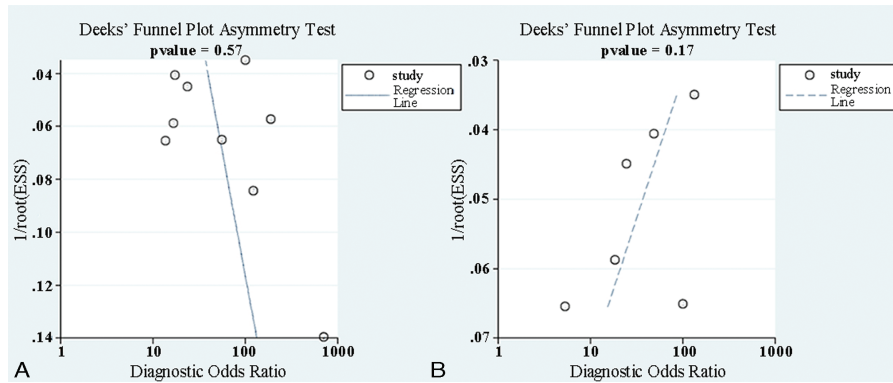
	Sensitivity	Specificity	Heterogeneity ( <i>P</i> value, <i>I</i> <sup>2</sup> )	
			Sensitivity	Specificity
Cut-off value of serum IgG4(135–144 mg/dL)				
Overall	0.872 (0.852–0.89)	0.826 (0.816–0.836)	0.000, 85.8%	0.000, 98%
Sample size				
Small	0.877 (0.839–0.909)	0.772 (0.744–0.799)	0.000, 91.3%	0.000, 96.2%
Large	0.870 (0.846–0.892)	0.837 (0.826–0.847)	0.000, 81.5%	0.000, 98.6%
Regions				
Asia	0.883 (0.862–0.903)	0.836 (0.825–0.847)	0.000, 86.2%	0.000, 98.3%
Europe and America	0.826 (0.773–0.872)	0.771 (0.742–0.798)	0.000, 86.4%	0.000, 97.5%
Diseases				
IgG4-RD	0.878 (0.852–0.901)	0.811 (0.799–0.822)	0.000, 85.9%	0.000, 98.5%
IgG4-associated cholangitis	0.864 (0.831–0.892)	0.888 (0.868–0.906)	0.000, 90.1%	0.01 78.2%
Cut-off value of serum IgG4(270–280 mg/dL)				
Overall	0.63 (0.60–0.66)	0.948 (0.941–0.954)	0.000, 91.3%	0.000, 86.5%
Sample size				
Small	0.483 (0.419–0.548)	0.949 (0.933–0.963)	0.000, 90.4%	0.000, 87.4%
Large	0.677 (0.643–0.711)	0.947 (0.939–0.954)	0.000, 74.6%	0.000, 90.5%
Regions				
Asia	0.677 (0.643–0.711)	0.947 (0.939–0.954)	0.02 74.6%	0.000, 90.5%
Europe and America	0.483 (0.419–0.548)	0.949 (0.933–0.963)	0.000, 90.4%	0.000, 87.4%
Diseases				
IgG4-RD	0.623 (0.578–0.667)	0.938 (0.929–0.946)	0.000, 94.2%	0.11 54.6%
IgG4-associated cholangitis	0.636 (0.593–0.678)	0.976 (0.966–0.985)	0.000, 91.2%	0.152 46.9%

IgG4-RD = Immunoglobulin G4-related disease.

**TABLE 3.** Meta-Analysis Estimates, Name Given to Study is Omitted

Study Omitted	Sensitivity (95% CI)	Specificity (95% CI)	Heterogeneity ( <i>P</i> value, <i>I</i> <sup>2</sup> )	
			Sensitivity	Specificity
Cut-off value of serum IgG4 (135–144 mg/dL)				
Carruthers 2015	0.87 (0.849–0.889)	0.839 (0.829–0.849)	0.000, 87.4%	0.000, 97.6%
Boonstra 2014	0.87 (0.849–0.889)	0.824 (0.814–0.834)	0.000, 87.4%	0.000, 98.2%
Oseini 2011	0.886 (0.866–0.904)	0.823 (0.813–0.834)	0.000, 81%	0.000, 98.2%
Yamamoto 2012	0.865 (0.844–0.884)	0.821 (0.811–0.832)	0.000, 85.6%	0.000, 98.1%
Li 2015	0.887 (0.866–0.906)	0.825 (0.814–0.835)	0.000, 85%	0.000, 98.2%
Yu 2015	0.877 (0.856–0.896)	0.879 (0.867–0.891)	0.000, 87.1%	0.000, 97.4%
Masaki 2012	0.86 (0.839–0.88)	0.826 (0.816–0.836)	0.000, 82.2%	0.000, 98.2%
Su 2015	0.871 (0.851–0.889)	0.798 (0.786–0.809)	0.000, 86.8%	0.000, 96.2%
Ohara 2013	0.862 (0.838–0.884)	0.817 (0.806–0.827)	0.000, 86.8%	0.000, 98%
Cut-off value of serum IgG4 (270–280 mg/dL)				
Carruthers 2015	0.652 (0.62–0.683)	0.95 (0.943–0.957)	0.000, 87.5%	0.000, 86%
Boonstra 2014	0.624 (0.592–0.656)	0.945 (0.938–0.952)	0.000, 92.8%	0.000, 86.5%
Oseini 2011	0.652 (0.62–0.684)	0.947 (0.939–0.953)	0.000, 89.6%	0.000, 88.8%
Li 2015	0.633 (0.597–0.668)	0.948 (0.941–0.955)	0.000, 93%	0.000, 89%
Yu 2015	0.606 (0.572–0.64)	0.958 (0.947–0.967)	0.000, 91%	0.000, 87.1%
Ohara 2013	0.602 (0.563–0.64)	0.943 (0.935–0.95)	0.000, 92.1%	0.000, 76.2%

CI = confidence interval.



**FIGURE 5.** Publication bias of serum IgG4 for IgG4-RD diagnosis. Funnel plot for the cut-off value of the normal upper limit (A) and 2-fold the normal upper limit (B). IgG4-RD = IgG4-related disease.

which can ensure the selection of patients with IgG4-RD and overall controls.

Several limitations in this meta-analysis need to be acknowledged. First, only a relatively small number of studies with limited subjects were included in the meta-analysis, which may reduce the statistical power for determining the diagnostic role of serum IgG4. Second, as mentioned above, significant heterogeneity was observed in this analysis. Because the number of included studies was <10, a multivariate meta-regression was not performed. New studies should be specifically designed in order to identify the possible confounding factors.

In conclusion, this study is the first to assess the diagnostic accuracy of serum IgG4 in patients with IgG4-RD using meta-analysis, which provided evidence that serum IgG4 is a modestly effective marker for the disease. Doubling the cut-off value for IgG4 could improve the specificity, but inevitably accompany with a significant sacrifice in sensitivity.

## REFERENCES

- Mahajan VS, Mattoo H, Deshpande V, et al. IgG4-related disease. *Annual Rev Pathol Mech Dis.* 2014;9:315–347.
- Culver EL, Vermeulen E, Makuch M, et al. Increased IgG4 responses to multiple food and animal antigens indicate a polyclonal expansion and differentiation of pre-existing B cells in IgG4-related disease. *Ann Rheum Dis.* 2015;74:944–947.
- Umehara H, Okazaki K, Masaki Y, et al. Comprehensive diagnostic criteria for IgG4-related disease (IgG4-RD), 2011. *Mod Rheumatol.* 2012;22:21–30.
- 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;2012:357071.
- Wallace ZS, Mattoo H, Carruthers M, et al. Plasmablasts as a biomarker for IgG4-related disease, independent of serum IgG4 concentrations. *Ann Rheum Dis.* 2015;74:190–195.
- Hamano H, Kawa S, Horiuchi A, et al. High serum IgG4 concentrations in patients with sclerosing pancreatitis. *N Engl J Med.* 2001;344:732–738.
- Carruthers MN, Khosroshahi A, Augustin T, et al. The diagnostic utility of serum IgG4 concentrations in IgG4-related disease. *Ann Rheum Dis.* 2015;74:14–18.
- Yamamoto M, Tabeya T, Naishiro 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;22:419–425.
- Oseini AM, Chaiteerakij R, Shire AM, et al. Utility of serum immunoglobulin G4 in distinguishing immunoglobulin G4-associated cholangitis from cholangiocarcinoma. *Hepatology.* 2011;54:940–948.
- Boonstra K, Culver EL, Wenniger LMD, et al. Serum immunoglobulin G4 and immunoglobulin G1 for distinguishing immunoglobulin G4-associated cholangitis from primary sclerosing cholangitis. *Hepatology.* 2014;59:1954–1963.
- Ohara H, Nakazawa T, Kawa S, et al. Establishment of a serum IgG4 cut-off value for the differential diagnosis of IgG4-related sclerosing cholangitis: a Japanese cohort. *J Gastroenterol Hepatol.* 2013;28:1247–1251.
- Li P, Chen H, Deng C, et al. Establishment of a serum IgG4 cut-off value for the differential diagnosis of IgG4-related disease in Chinese population. *Mod Rheumatol.* 2015:1–5.
- Yu KH, Chan TM, Tsai PH, et al. Diagnostic performance of serum IgG4 levels in patients with IgG4-related disease. *Medicine.* 2015;94:e1707.
- Su YY, Sun W, Wang CQ, et al. Detection of serum IgG4 levels in patients with IgG4-related disease and other disorders. *Plos One.* 2015;10:e0124233.
- Masaki Y, Kurose N, Yamamoto M, et al. Cutoff values of serum IgG4 and histopathological IgG4+ plasma cells for diagnosis of patients with IgG4-related disease. *Int J Rheumatol.* 2012;2012:580814.
- Morselli-Labate AM, Pezzilli R. Usefulness of serum IgG4 in the diagnosis and follow up of autoimmune pancreatitis: a systematic literature review and meta-analysis. *J Gastroenterol Hepatol.* 2009;24:15–36.
- Lian MJ, Liu S, Wu GY, et al. Serum IgG4 and IgG for the diagnosis of autoimmune pancreatitis: a systematic review with meta-analysis. *Clin Res Hepatol Gastroenterol.* 2016;40:99–109.
- Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Int Med.* 2011;155:529–536.
- Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. *BMJ.* 2003;327:557–560.
- Glas AS, Lijmer JG, Prins MH, et al. The diagnostic odds ratio: a single indicator of test performance. *J Clin Epidemiol.* 2003;56:1129–1135.
- Shimosegawa T, Chari ST, Frulloni L, et al. International consensus diagnostic criteria for autoimmune pancreatitis: guidelines of the International Association of Pancreatology. *Pancreas.* 2011;40:352–358.
- Umehara H, Okazaki K, Masaki Y, et al. A novel clinical entity, IgG4-related disease (IgG4RD): general concept and details. *Mod Rheumatol.* 2012;22:1–14.