

Comparison of Commercially Available Serologic Kits for the Detection of Celiac Disease

Afzal J. Naiyer, MD, Lincoln Hernandez, MD, Edward J. Ciaccio, PhD,
Konstantinos Papadakis, MD, John S. Manavalan, MD, Govind Bhagat, MD,
and Peter H. R. Green, MD

Background: The sensitivity and specificity of current antihuman tissue transglutaminase (tTG) IgA assays used to detect celiac disease reportedly approach 100%. In addition, the sensitivity of new generation deamidated gliadin peptide (α -DGP) antibody assays has also been reported to be similar to the tTG IgA assays. In routine clinical practice, however, the sensitivities and specificities of these tests for diagnosing celiac disease seem to be lower.

Aim: We analyzed sensitivities and specificities of 4 IgA tTG and 3 deamidated gliadin peptide (α -DGP) kits.

Methods: The performance of 4 tTG IgA assays, A: Inova (Hu red blood cell), B: Binding site (rHu Ag), C: Eurospital (rHu Ag), D: Immco (rHu Ag) and 3 Inova α -DGP assays, E: α -DGP-IgA, F: α -DGP-IgG, and G: α -DGP-IgA + G was evaluated using sera from different subsets of celiac disease patients and controls; group 1: active celiac disease n = 28, group 2: gluten-free diet n = 54, group 3: healthy controls n = 40, group 4: disease controls n = 57 (Crohn's disease n = 17, chronic hepatitis n = 40).

Results: Using the manufacturer's cut-off values, the sensitivities and specificities of different kits ranged from 71.4% to 96.4% and 87.5% to 100%, respectively. When group 1 was compared with disease controls, sensitivities remained the same but specificities decreased. Receiver operating characteristic plot derived cut-off values modified decision thresholds in all assays except kit (G). Kappa analysis demonstrated variable degrees of agreement. All assays demonstrated higher sensitivities for patients with higher grades of villous atrophy.

Conclusions: Overall sensitivity was at or below 90%, which is lower than that reported in the literature. Performance of the recombinant and red blood cell antigen-based tTG assays was similar, whereas the α -DGP assays demonstrated lower values. Receiver operating characteristic plot derived cut-off values altered test results. Many factors affect the results of these tests and clinicians should be aware of their limitations.

Key Words: tTG, ELISA, celiac disease

(*J Clin Gastroenterol* 2009;43:225–232)

Celiac disease is an autoimmune like inflammatory disease of the small intestine resulting from gluten ingestion by genetically susceptible individuals.¹ It occurs

worldwide with an estimated frequency of approximately 1%.² At present, the detection of human tissue transglutaminase (tTG) antibodies using human recombinant or red cell tTG-based enzyme-linked immunosorbent blood assays (ELISAs) and antiendomysial antibodies by immunofluorescence techniques are considered equivalent in detecting celiac disease.^{3,4} The ELISA-based assays allow for automation and as a result are more widely used in clinical practice. A review commissioned by the National Institutes of Health (NIH) that led to the Consensus Development Conference on Celiac Disease in June 2004, revealed the pooled estimates of sensitivities and specificities for antihuman tTG IgA assays, in adults, to be 98.1% and 98.0%, respectively.³

Several studies have questioned the sensitivity of tTG assays in the clinical practice setting.^{4–6} The issue of the sensitivity of human antigen-based tTG assays in diagnosing celiac disease is further compounded by the observation of false-positive tTG IgA tests in patients with a variety of diseases, including chronic liver disease,^{7,8} end-stage heart failure,⁹ diabetes,¹⁰ inflammatory bowel disease,^{8,11} and arthritis.¹² In addition, studies comparing different human tTG test kits, from different manufacturers, have revealed variable sensitivities for detecting celiac disease.^{13–15}

We therefore compared 4 frequently available human IgA tTG ELISA-based assays that are currently used in the United States and 3 new generation antigliadin antibody assays that have been developed on the principle that gliadin reactive antibodies from celiac disease patients recognize tTG deamidated nanopptide epitopes of gliadin. These antibodies are specific and sensitive for celiac disease and dermatitis herpetiformis, equivalent to recombinant tTG assays. The old generation antigliadin antibodies are considered to be nonspecific occurring in inflammatory bowel disease,¹⁶ IgA nephropathy,¹⁷ HIV infections,¹⁸ neurologic disorders,¹⁹ and rheumatoid arthritis.²⁰ We used these tTG and DGP assays to determine their diagnostic accuracy and their degree of agreement in patients with celiac disease.

MATERIALS AND METHODS

Study Population

We used stored, frozen serum samples that were obtained after informed consent from patients with celiac disease or patients who were being evaluated for celiac disease at the Celiac Disease Center of Columbia University. Study was approved by the Columbia University Institutional Review Board.

Group 1 (n = 28) consisted of consecutive adult patients with biopsy-confirmed active celiac disease (ACD). Blood samples were obtained at the time of first biopsy.

Received for publication June 12, 2007; accepted November 10, 2007. From the Department of Medicine and Pathology, Columbia University, New York, NY.

No conflicts of interest exist.

This study has been self funded by us, The Celiac Disease Center at Columbia University.

Reprints: Dr Peter H. R. Green, MD, Department of Medicine, College of Physicians and Surgeons, Columbia University, 180 Ft Washington Av (Room 956), New York, NY 10032 (e-mail: pgl1@columbia.edu). Copyright © 2009 by Lippincott Williams & Wilkins

Group 2 (n = 54) consisted of celiac disease patients on a gluten-free diet (GFD). The duration of GFD in group 2 ranged from 2 to 87 months. Dietary adherence was determined by both the celiac center nutritionist and a physician. At the time of blood drawing 44 patients were considered to be strictly adherent to the diet. Some patients (n = 10) were considered to be initially poorly compliant with the diet. These 10 patients were therefore not included in the analysis of antibody titers and duration on the diet.

Group 3 (n = 40) consisted of healthy adults who served as normal controls. These subjects were recruited from the University and laboratory staff, gave consent and upon specific questioning denied gastrointestinal symptoms.

Group 4 (n = 57) consisted of adult patients with other diseases including Crohn's disease (n = 17; serum obtained from Cedars-Sinai Medical Center, Los Angeles, CA) and chronic hepatitis C virus infection (n = 40; serum obtained from patients enrolled in a prospective study at the Department of Gastroenterology, Columbia University).

Antibody Testing

Four commercially available tTG IgA ELISA kits:

- (a) Inova [Hu red blood cell (RBC) tTG IgA].
- (b) Binding site (recombinant human Ag).
- (c) Eurospital (recombinant human Ag).
- (d) Immco (recombinant human Ag).

And 3 new generation deamidated gluten peptide kits (α-DGP) ELISA kits.

- (e) Inova [Antigliadin antibody (AGA) II-IgA].
- (f) Inova (AGA II-IgG).
- (g) Inova (AGA-II IgA + IgG), a conjugate kit.

All kits were used according to the manufacturer's protocol. Binding site (Birmingham, UK), Eurospital (Trieste, Italy), Immco (Buffalo, NY) Inova (San Diego, CA). The details of the different assays are summarized in Table 1. All assays were performed manually and the optical density was measured with Multiskan EX (Thermo Electron Corporation, Vantaa, Finland). The manufacturer recommended cut-off values and cut-off values determined by receiver operating characteristic (ROC) plots were used to calculate the diagnostic performance of each test.

Small Bowel Histology

Biopsy specimens were obtained from the second duodenal portion during gastroduodenoscopy. Histologic evaluation was performed according to the modified Marsh classification as described by Oberhuber.²¹ We only included in the analysis biopsy results of those patients in whom we were able to obtain and review the pathologic materials (n = 26).

Statistical Analysis

The diagnostic performance of each assay or the ability to discriminate diseased from normal individuals was evaluated using ROC curve analysis, thus selecting the cut-offs that provided the best combination of sensitivity and specificity. Areas under the ROC curves and their 95% confidence intervals were calculated for each assay using the nonparametric method described by Hanley and McNeil²² and compared using the methodology of DeLong et al.²³ The latter adjusts for possible correlations arising due to the fact that the same individuals underwent all tests for

TABLE 1. Assay Characteristics

Manufacturer	tTG Antigen	Serum Dilution	Incubation Times	No. Calibrators	Conjugate	Substrate	Proposed Cut-off	Measuring Range	ROC Cut-off
A. Inova	Purified human/erythrocyte	1:100	30;30;30	1	Goat HRP/ antihuman IgA	TMB	20 U	0 U	21
B. Binding site	Recombinant human/ <i>Escherichia coli</i>	1:100	30;30;30	5	Purified Peroxidase/ antihuman IgA	TMB	4 kU/L	1.23-100 kU/L	6
C. Eurospital	Recombinant human/ <i>Escherichia coli</i>	1:25	60;60;30	1	Goat HRP/ antihuman IgA	TMB	7 kU/L	0 kU/L	12
D. Immco	Recombinant human/ <i>Escherichia coli</i>	1:50	60;30;30	4	AP/antihuman IgA antihuman IgA	pNPP	20 kU/L	11-179 kU/L	2
E. Inova-AGA-II (α-DGP) IgA	Synthetic gliadin-derived deamidated peptides	1:100	30;30;30	1	Goat HRP/ antihuman IgA	TMB	20 U	0 U	16
F. Inova-AGA-II (α-DGP) IgG	Synthetic gliadin-derived deamidated peptides	1:100	30;30;30	1	Goat HRP/ antihuman IgA	TMB	20 U	0 U	7
G. Inova-AGA-II (α-DGP) IgA + IgG	Synthetic gliadin-derived deamidated peptides	1:100	30;30;30	1	Goat HRP/ antihuman IgA	TMB	20 U	0 U	18

HRP indicates horseradish peroxidase; pNPP, paramitrophenyl phosphate; TMB, 3,3',5,5', tetramethylbenzidine.

which area under the ROC curve (AUROC) was calculated. Bland Altman analysis, Pearson correlation coefficient, and Demming regression were also performed to observe the degree of agreement between assays. ROC plot analysis, linear correlation, Demming regression and Altman-Bland analysis were performed with MedCalc Software (MedCalc v. 7.2.1.0, Mariakerke, Belgium). Sigma Plot (Systat Software Inc, ver 10) was also used for the statistical analysis.

RESULTS

Diagnostic Performance of tTG IgA Assays

Sensitivity and Specificity ACD Patients Compared With Normal Controls

Sensitivity and specificity were calculated for each method using the manufacturers and ROC plot derived cut-off values (Table 2). Using the manufacturer recommended cut-off values, sensitivities, and specificities of the tTG IgA kits varied from 85.7% to 96.4% and 87.5% to 100%, respectively.

Adjusting the cut-off threshold according to the ROC plot analyses (with the highest sum of sensitivity and specificity) altered the diagnostic accuracy of all tTG IgA kits. The effects ranged from increasing the sensitivities and decreasing the specificities in 2 kits (A and B) to unchanged sensitivity and increased specificity for kit D. However, 1 kit (kit C) demonstrated increased sensitivity and unchanged specificity after the employment of ROC plot derived cut-off. This kit also had the highest sum of sensitivity and specificity before and after incorporating the ROC plot derived cut-off value among all tTG assays (Table 2). On the basis of the ROC plot analyses, the areas under the curves were 0.962, 0.976, 0.979, and 0.972, for assays A to D, respectively, with no statistically significant difference between the tests.

Sensitivity and Specificity ACD Patients Compared With Disease Controls

Using the manufacturer's cut-off values, overall sensitivities obtained for all tTG IgA kits (Table 2) were similar to those obtained when patients with ACD were compared with normal controls, except kit C that showed increased sensitivity. Specificities of all kits decreased with the exception of 1 kit (kit A), which showed increased specificity. Again, kit C demonstrated the highest sum of sensitivity and specificity in this comparison group (Table 2).

ROC curve analyses altered the sensitivities and specificities of the different test kits. Two kits demonstrated increased sensitivities (A, B) and 2 unchanged sensitivities (C and D). The specificity of 1 kit was increased (D), unchanged in 2 (A and C), and decreased in 1 (B) (Table 2). Again, kit C had the highest sum of sensitivity and specificity after the ROC derived cut-off was used (Table 2). The areas under the curves for kits A to D were 0.966, 0.961, 0.974, and 0.955, respectively, not statistically different among the kits.

Diagnostic Performance of the New Generation Antigliadin (α-DGP) Assays

Sensitivity and Specificity ACD Patients Compared With the Normal Controls

The α-DGP assays (kits E, F, and G) demonstrated sensitivities and specificities ranging from 71.4% to 82.1%

TABLE 2. Sensitivity and Specificity of Celiac Disease Diagnostic Tests

	A	B	C	D	E	F	G
	Inova tTG-IgA	Binding Site tTG-IgA	Eurospiral tTG-IgA	Immco tTG-IgA	Inova DGP-IgA	Inova DGP-IgG	Inova DGP-Ig(A+G)
A. Active celiac disease vs. normal control							
Manufacturers cut-off	20 kU	4 kU	7 kU	20 kU	20 kU	20 kU	20 kU
Sensitivity (%)	89.3	85.7	92.9	96.4	78.6	71.4	82.1
Specificity (%)	95	100	100	87.5	100	95	97.5
Cut-off on the basis of the ROC	12 kU	2 kU	6 kU	21 kU	16 kU	7 kU	18 kU
Sensitivity (%)	92.9	96.4	96.4	96.4	85.7	92.9	82.1
Specificity (%)	95	95	100	95	97.5	92.5	97.5
B. Active celiac disease vs. disease control							
Manufacturers cut-off	20 kU	4 kU	7 kU	20 kU	20 kU	20 kU	20 kU
Sensitivity (%)	89.3	85.7	96.4	96.4	78.6	71.4	82.1
Specificity (%)	98.3	93.1	98.3	75.9	94.8	98.3	98.3
Cut-off on the basis of the ROC	14 kU	2 kU	6 kU	22 kU	16 kU	7 kU	20 kU
Sensitivity (%)	92.9	96.4	96.4	96.4	85.7	92.9	82.1
Specificity (%)	98.3	87.9	98.3	89.7	93.1	84.5	98.3

Results greater than or equal to cut-off threshold were considered positive both with the manufacturers and ROC analysis derived values. Specificity was calculated using normal controls and disease controls.

and 95% to 100%, respectively (Table 2). Adjusting the cut-off threshold according to the ROC plot analyses altered the diagnostic accuracy of kits (E and F) with increasing sensitivity and decreasing specificity, whereas kit G demonstrated similar sensitivity and specificity before and after the employment of ROC plot derived cut-off (Table 2). The areas under the curves were 0.903, 0.951, and 0.926 for kits E, F, and G, respectively, and were not statistically significantly different among the AGA-II kits.

Sensitivity and Specificity Active Celiac Patients Compared With Disease Controls

The AGA assays demonstrated sensitivities (Table 2) similar to the ones observed when patients with ACD were compared with normal controls (Table 2), however, the specificities of kits F and G increased but the specificity of kit E decreased. ROC plot analyses altered the diagnostic accuracy by increasing sensitivity and decreasing specificity of kits (E and F), sensitivity and specificity of kit G was not altered. The areas under the curves were 0.879, 0.942, and 0.923 for kits E, F, and G, respectively, not statistically different among the kits. However, a statistically significant difference ($P < 0.05$) was observed between the areas under the curve of kits A and E and kits C and E.

Rate of False-positive Tests in the tTG-IgA and α -DGP Assays

The rate of false-positive tTG IgA and AGA-II (α -DGP) assays when patients with ACD were compared with normal controls is shown in Table 3. With the manufacturer's cut-off values all kits demonstrated a relatively low false-positive rate (0% to 8%) particularly kit C which showed a false-positive rate of 0% in the control groups. Kit D, however, showed a higher false-positive rate, 13% in normal controls, 24% in patients with Crohn's disease, and 25% in patients with chronic hepatitis C, which decreased after the employment of the ROC plot derived cut-off. The decrease in the false-positive rate of kit D after ROC plot derived cut-off parallels its increased specificity after ROC curve analysis. The false-positive rates were higher in the disease control comparison.

Method Comparison

The assay characteristics and the manufacturer's instructions (Table 1) were documented for all assays. Using κ analysis we assessed agreement among the different

assays. The spectrum of the strength of agreement was interpreted as follows: < 0.2 poor, 0.21 to 0.40 fair, 0.41 to 0.60 moderate, 0.61 to 0.80 good, and 0.81 to 1.00 for a very good agreement. The κ score ranged between 0.49 and 0.819 for comparison between all assays.

Between tTG Assays

Kappa analysis demonstrated a good agreement between all tTG IgA kits (κ ; 0.61 to 0.80). Two kits (A and C) demonstrated very good agreement (κ ; 0.819), whereas 2 kits (B and D) had a moderate agreement (κ ; 0.554).

Between α -DGP Assays

The combined isotype kit G showed a good agreement with the kits E and F individually, however, a moderate agreement (κ ; 0.592) was observed between the kits E and F.

Between tTG and α -DGP Assays

Kappa analysis showed a good agreement between the kit E and the tTG IgA assays except with kit (D), where it showed a moderate agreement. Kit F and the combined isotype kit G, however, showed a moderate agreement with the tTG IgA assays.

Overall 2 kits (A and C) showed the best agreement with all other tests. Linear correlation between assays gave similar results to κ analysis.

Correlation Between Assay Titers and Degree of Villous Atrophy in Patients With ACD

We assessed the sensitivity of all tTG IgA and α -DGP assays in detecting the degree of villous atrophy before and after employment of ROC curve derived cut-offs for analysis (Table 4). Sensitivity for detecting the degree of villous atrophy varied among the kits. The new generation α -DGP assays were less sensitive (62% to 69%) in detecting mild mucosal damage that is, partial villous atrophy compared with the tTG IgA (85% to 92%) assays. ROC analysis derived cut-offs resulted in increased sensitivities of kits B (69% to 85%), E (69% to 85%), and F (62% to 85%), however, the sensitivities of all other assays remained unchanged.

For the more severe degree of villous atrophy (subtotal villous atrophy and total villous atrophy), higher sensitivities, 92% for kits A and B and 100% for kits C and D, were observed. The α -DGP kits F and G demonstrated

TABLE 3. Rate of False-positive Tests in Control Groups Using Cut-off Values From Celiac Disease Versus Normal Control Comparison

Test	Manufacturer's-Cut-off Point				ROC Plot Analysis-Cut-off Point			
	Cut-off (kU)	Normals (%)	Hepatitis (%)	Crohn's (%)	Cut-off (kU)	Normals (%)	Hepatitis (%)	Crohn's (%)
Inova tTG IgA (A)	20	2/40 (5)	1/40 (3)	0/17 (0)	12	2/40 (5)	2/40 (5)	0/17 (0)
Binding site tTG IgA (B)	4	0/40 (0)	3/40 (8)	0/17 (0)	2	7/40 (18)	12/40 (30)	2/17 (12)
Eurospital tTG IgA (C)	7	0/40 (0)	0/40 (0)	0/17 (0)	6	2/40 (5)	2/40 (5)	0/17 (0)
Immco tTG IgA (D)	20	5/40 (13)	10/40 (25)	4/17 (24)	21	4/40 (10)	10/40 (25)	3/17 (18)
Inova anti-DGP IgA (E)	20	0/40 (0)	2/40 (5)	1/17 (6)	16	2/40 (5)	3/40 (8)	2/17 (12)
Inova anti-DGP IgG (F)	20	2/40 (5)	0/40 (0)	1/17 (6)	7	5/40 (13)	6/40 (15)	6/17 (35)
Inova anti-DGP IgA + IgG (G)	20	1/40 (3)	2/40 (5)	0/17 (0)	18	2/40 (5)	3/40 (8)	1/17 (6)

Results equal to or greater than the cut-off were considered positive both with the manufacturers and ROC analysis derived values. Specificity was calculated using normal controls and non-celiac disease controls. Crohn's indicates Crohn's disease; hepatitis, chronic HCV; normals, healthy control.

TABLE 4. Sensitivity of Assays in Detecting the Degree of Villous Atrophy

Test	Test Result	Manufacturers Cut-off				ROC Cut-off			
		PVA		TVA		PVA		TVA	
		n = 13	Sensitivity (%)	n = 13	Sensitivity (%)	n = 13	Sensitivity (%)	n = 13	Sensitivity (%)
Inova tTG IgA (A)	Positive	12	92	12	92	12	92	13	100
	Negative	1		1		1		0	
Binding site tTG IgA (B)	Positive	11	85	12	92	12	92	13	100
	Negative	2		1		1		0	
Eurospital tTG IgA (C)	Positive	12	92	13	100	12	92	13	100
	Negative	1		0		1		0	
Immco tTG IgA (D)	Positive	12	92	13	100	12	92	13	100
	Negative	1		0		1		0	
Inova anti-DGP IgA (E)	Positive	9	69	13	100	11	85	13	100
	Negative	4		0		2		0	
Inova anti-DGP IgG (F)	Positive	8	62	10	77	11	85	13	100
	Negative	5		3		2		0	
Inova anti-DGP IgA + IgG (G)	Positive	9	69	12	92	9	69	12	92
	Negative	4		1		4		1	

ROC cut-off indicates receiver operator curve cut-off.

lower sensitivities than the tTG assays. ROC derived cut-offs increased the sensitivities to 100% for all assays with the exception of kit G.

We did not observe any correlation between the assay titers and the degree of villous atrophy in patients with ACD. A trend of worsening villous atrophy with increasing antibody titer levels was observed, however, it was not statistically significant except kit E in which, significance was observed between the assay titers in partial and total villous atrophy.

Antibody Profile in Celiac Disease Patients on a GFD

To determine the sensitivity of the 7 assays for detection of adequate GFD adherence, we compared the percent of positive patients on a GFD for the tTG and α -DGP assays. When the titers were correlated with the duration of GFD we observed a linear correlation between increasing time in months on GFD and decreasing antibody titers. A statistically significant correlation was observed in 4 assays (A, D, E, F, and G) but not with kits B and C (Fig. 1). From our

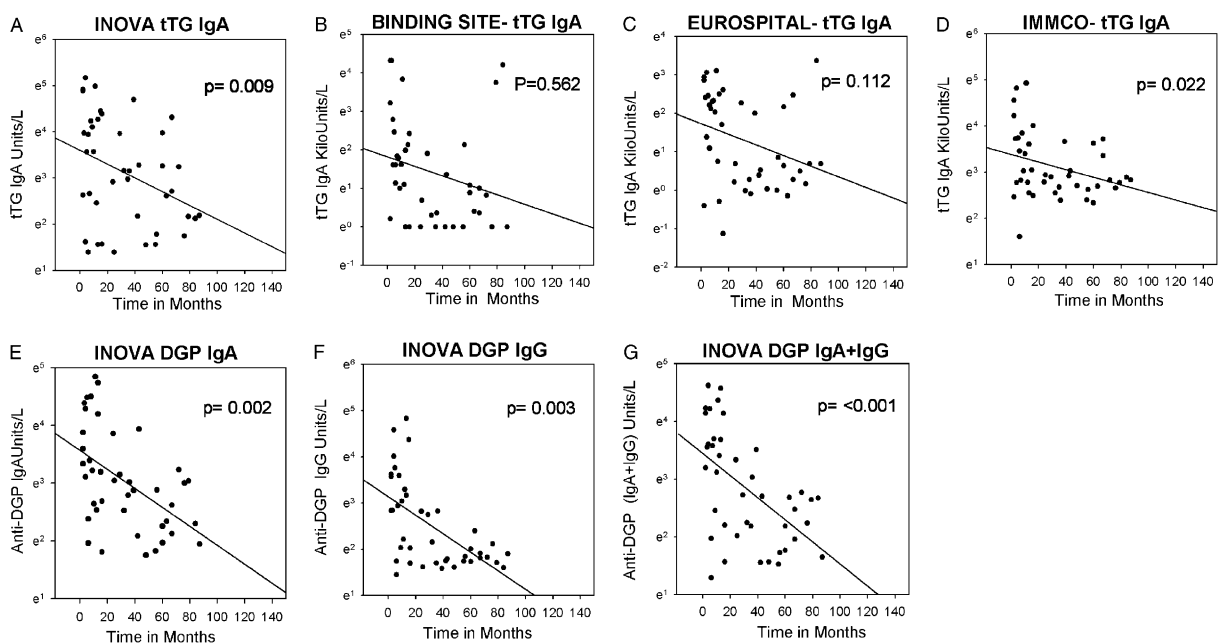


FIGURE 1. Linear correlation between assay titers and time on GFD. X-axis is plotted with time in months on GFD. Antibody titer in log phase for each respective assay was plotted on the Y-axis. All assays except B and C demonstrated a statistical significance for the linear correlation between increasing time in months on a GFD and antibody titers. P value of <0.05 was considered significant. A, D, E, F, and G denote the letter assigned to the kits as in Table 1.

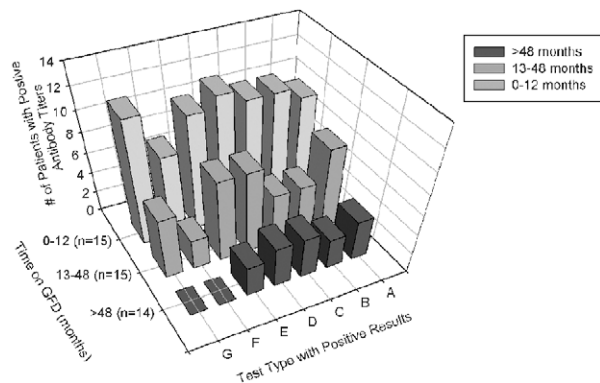


FIGURE 2. Assay profile of patients on a GFD. X-axis shows test types, A to D (tTG IgA assays) E to G (DGP assays), Y-axis is the number of patients with positive test results and Z-axis denotes the 3 time points on the GFD along with the number of patients (n) in each time point.

data it seems that kit G (DGP IgA + G) was the best test to detect compliance in patients on a GFD, which had a *P* value of < 0.001 for the correlation between time on diet and assay titer (Fig. 1). The GFD group was divided into 3 categories depending on the duration on diet, < 12 months, 13 to 48 months, and > 48 months on diet groups. Figure 2 shows the number of positive results for each assay according to duration on the GFD. These results show that the IgG DGP-based assays (F and G) seem to have greater sensitivity for detection of adequate GFD adherence.

DISCUSSION

In view of the studies demonstrating disappointing sensitivities and specificities of tTG IgA assays in the clinical practice setting,^{4,6,13-15,24} we tested the diagnostic accuracy of 4 commercially available tTG IgA assays frequently used in the United States. In our study, the sensitivity of the tTG tests did not reach that reported in a review of the literature, which was undertaken for the NIH Consensus Development Conference on Celiac Disease in 2004.^{2,3,25}

In addition, our study demonstrates that the sensitivity of the different tTG IgA kits varied from 85.7% to 96.4%. One Kit (D, Immco) a human recombinant tTG-based assay had the best sensitivity of 96.4% with the manufacturer's recommended cut-off value. However, this kit showed poor specificity (87.5%) at this level. Of interest, the source of assay antigen, either human recombinant or human RBCs, did not seem to be a major factor in determining the sensitivity of the tTG IgA assays.

To estimate the accuracy of the tests in our population, we performed ROC curve analysis by comparing the different celiac disease patient groups with both, healthy and disease control groups. When the ROC-derived cut-off values were employed the sensitivities of all kits increased, except 1 Kit (D, Immco) where it remained unchanged. Although the sensitivities of tests for antihuman tTG IgA have been reported in the literature as being approximately 98%,³ our data aligns well with the sensitivities obtained in the clinical setting $\leq 90\%$.^{4,5,26} When we looked at the false-positive rates of the kits we noted an increase in the false-positive rates especially in the Crohn's Disease and Hepatitis control groups after the employment of the ROC-

derived cut-off values. The change in results using ROC curve analysis suggests that this is another variable that both laboratories and clinicians should take into account. Although ROC curve analysis is frequently used in the research setting, testing laboratories usually use the manufacturer's determined cut-off values for determining results.

It is of interest to determine factors that contribute to the sensitivity of these tests. The degree of villous atrophy is important in this respect. We observed all tTG assays to have 100% sensitivity for total villous atrophy and 92% for partial villous atrophy. Our observed assay sensitivity for total villous atrophy and partial villous atrophy is higher than that reported previously in studies^{5,27} using guinea pig tTG IgA kits. Our study did not demonstrate a correlation between antibody titers and the degree of villous atrophy, interestingly, when the mean tTG IgA antibody titers from each assay was plotted against the degree of villous atrophy we only observed a trend of increasing antibody titers with increasing degree of villous atrophy. The values were not significant.

The specificity of the tTG IgA kits ranged from 87.5% to 100%, however, when the celiac disease group was compared with disease controls the specificities of all kits were reduced, with further reduction after the employment of ROC-derived cut-off value with the exception of 1 kit (A, Inova). The specificities obtained in our study are less than those observed by Rostom et al³ (99%). This is an important consideration because in clinical practice the patient population that is being tested will contain not only patients with celiac disease but also other diseases such as Crohn's disease, rather than healthy individuals. Our data suggests that false-positive rates are kit specific and are further increased after employing ROC curve-derived thresholds.

Old generation antigliadin assays for detecting antibodies against gliadin are usually considered much less specific and sensitive for celiac disease than autoantibodies²⁸ as antigliadin antibodies have been reported in a variety of conditions not related to celiac disease. Recent work has revealed that deamidation of gliadin by tTG, results in enhanced binding by antigliadin antibodies. Hence the new generation antigliadin (DGP) assays using the deamidated peptides have been shown to have comparable diagnostic accuracy for celiac disease when compared with human recombinant or RBC antigen-based anti-tTG antibody detection assays.⁹

Using these new generation antigliadin antibody assays we observed a significantly lower sensitivity and specificity of the DGP. Our data are in contrast to 2 other studies which have demonstrated a higher sensitivity of the DGP assays compared with tTG IgA assays.^{29,30} In the prospective study by Sugai et al²⁹ the initial diagnosis was made based only on endoscopy and when the biopsy and serologies (performed later) were compared with patient diagnosis, the sensitivity of DGP assays was demonstrated to be only marginally better than the tTG IgA assay. One possible explanation for our study showing a lower sensitivity for DGP assays may be due to a selection bias. Twenty-three of twenty-eight patients with ACD were initially diagnosed based on a positive tTG test, indicating that our celiac disease group could be enriched in tTG IgA-positive cases. Similar to the tTG IgA assays, a trend between worsening villous atrophy and increasing antibody titers was observed.

To determine the serologic tests that are sensitive for detection of adequate GFD adherence, we compared the sensitivity of the tTG IgA and DGP assays in our GFD population. Although tTG antibody titers have been shown to take approximately 12 months to decrease after introduction of a GFD,³¹ the level of α -DGP titers and duration of diet has not been evaluated. We observed that the rates of positive results using the tTG IgA and α -DGP assays were equivalent in the 0 to 12 and the 13 to 48 month group of our GFD patients. Interestingly, our data shows that among the deamidated gluten peptide assays the frequency of positive tests decreased significantly at the > 48 months period whereas the positive rate for tTG assays still ranged between 29% and 57%. To this end, we performed regression analysis to determine the overall performance of the assays in detecting compliance in the GFD population. As demonstrated in Figure 1 a statistically significant linear correlation was observed between increasing duration of GFD and decreasing test values. This indicates that removing gliadin from the diet leads to a more dramatic drop in gliadin peptide antibodies than tTG-antibodies, indicating that DGP IgA-based tests seem to have greater sensitivity for the detection of adequate GFD adherence.

To assess the degree of agreement between assays both κ and correlation analysis demonstrated that 2 kits (kits A and C) had the highest degree of agreement among assays. As a measure of diagnostic accuracy, demonstrated by AUROC, all the assays showed values of more than 0.90 indicating a superior test performance. Although both the tTG and DGP assays did not show any statistically significant difference in the AUC values, they demonstrated disparate sensitivities and specificities. Although the ROC-derived cut-off values change between tests and populations, it is not grounds for changing the cut-off value for the tests. The fact that the actual predictive values change from one test to another should be interpreted as just one piece of the diagnostic paradigm.

It is also important to note that none of the assays, either tTG or DGP, had a sensitivity of 100%. In the ACD group, tTG IgA kits demonstrated higher sensitivities than the α -DGP assays, however, we observed 3 individuals who were falsely negative by one of the 4 tTG IgA assays to be positive by one of the α -DGP assays, inversely 3 patients who were falsely negative by the α -DGP assays were detected by the tTG assays. These data indicate that one single test does not detect all those individuals with celiac disease, and that a more sensitive screening method could be achieved by employing tTG IgA and α -DGP assays in parallel or in combination.

In conclusion, our study demonstrated that the overall sensitivities of human recombinant or RBC tTG IgA assays is at or below 90%, which is lower than that reported in the literature. Although our results are produced by sample sizes that are significantly smaller than those by which manufacturers produced the cut-off values, all kits were tested with the same population and demonstrated variable sensitivities and specificities. Only 1 kit (C) approached what was considered optimal by the NIH consensus conference.² On comparing the tTG assays with the new generation antigliadin antibody (α -DGP) assays, we observed a lower sensitivity with the latter possibility owing to a selection bias. However, it may be of value to use these newer assays in combination with the α -tTG assays to increase the sensitivity for detecting celiac disease. The

α -DGP assays that detect antibodies to the deamidated gliadin peptide seem to have greater sensitivity for the detection of adequate GFD adherence. Our data also indicate that studies comparing antibody assays should compare the study population not only with normal controls but disease controls as well. Taking into account the variables in each kit and the inability of one test to be used interchangeably with another there is a need for standardization of all commercial assays in the United States as has been attempted in Europe.²⁸ Finally, it is important that clinicians be aware of the limitations of these tests.

REFERENCES

- Green PH, Jabri B. Coeliac disease. *Lancet*. 2003;362:383–391.
- National Institutes of Health. National Institutes of Health Consensus Development Conference Statement on Celiac Disease, June 28–30, 2004. *Gastroenterology*. 2005;128:S1–S9.
- Rostom A, Dube C, Cranney A, et al. The diagnostic accuracy of serologic tests for celiac disease: a systematic review. *Gastroenterology*. 2005;128:S38–S46.
- Abrams JA, Brar P, Diamond B, et al. Utility in clinical practice of immunoglobulin a anti-tissue transglutaminase antibody for the diagnosis of celiac disease. *Clin Gastroenterol Hepatol*. 2006;4:726–730.
- Dickey W, McMillan SA, Hughes DF. Sensitivity of serum tissue transglutaminase antibodies for endomysial antibody positive and negative coeliac disease. *Scand J Gastroenterol*. 2001;36:511–514.
- Fabiani E, Catassi C. The serum IgA class anti-tissue transglutaminase antibodies in the diagnosis and follow up of coeliac disease. Results of an international multi-centre study. International Working Group on Eu-tTG. *Eur J Gastroenterol Hepatol*. 2001;13:659–665.
- Vecchi M, Folli C, Donato MF, et al. High rate of positive anti-tissue transglutaminase antibodies in chronic liver disease. Role of liver decompensation and of the antigen source. *Scand J Gastroenterol*. 2003;38:50–54.
- Bizzaro N, Villalta D, Tonutti E, et al. IgA and IgG tissue transglutaminase antibody prevalence and clinical significance in connective tissue diseases, inflammatory bowel disease, and primary biliary cirrhosis. *Dig Dis Sci*. 2003;48:2360–2365.
- Peracchi M, Trovato C, Longhi M, et al. Tissue transglutaminase antibodies in patients with end-stage heart failure. *Am J Gastroenterol*. 2002;97:2850–2854.
- Leon F, Camarero C, R-Pena R, et al. Anti-transglutaminase IgA ELISA: clinical potential and drawbacks in celiac disease diagnosis. *Scand J Gastroenterol*. 2001;36:849–853.
- Di Tola M, Sabbatella L, Anania MC, et al. Anti-tissue transglutaminase antibodies in inflammatory bowel disease: new evidence. *Clin Chem Lab Med*. 2004;42:1092–1097.
- Spadaro A, Sorgi ML, Scivo R, et al. Anti-tissue transglutaminase antibodies in inflammatory and degenerative arthropathies. *Reumatismo*. 2002;54:344–350.
- Wong RC, Wilson RJ, Steele RH, et al. A comparison of 13 guinea pig and human anti-tissue transglutaminase antibody ELISA kits. *J Clin Pathol*. 2002;55:488–494.
- Fernandez E, Riestra S, Rodrigo L, et al. Comparison of six human anti-transglutaminase ELISA-tests in the diagnosis of celiac disease in the Saharawi population. *World J Gastroenterol*. 2005;11:3762–3766.
- Van Meensel B, Hiele M, Hoffman I, et al. Diagnostic accuracy of ten second-generation (human) tissue transglutaminase antibody assays in celiac disease. *Clin Chem*. 2004;50:2125–2135.
- Kull K, Uibo O, Salupere R, et al. High frequency of antigliadin antibodies and absence of antireticulin and antiendomysium antibodies in patients with ulcerative colitis. *J Gastroenterol*. 1999;34:61–65.

17. Ots M, Uibo O, Metskula K, et al. IgA-antigliadin antibodies in patients with IgA nephropathy: the secondary phenomenon? *Am J Nephrol*. 1999;19:453–458.
18. Zuin G, Fontana M, Morelli A, et al. Antigliadin antibodies in HIV-infected children. *Pediatr AIDS HIV Infect*. 1996;7:409–412.
19. Burk K, Bosch S, Muller CA, et al. Sporadic cerebellar ataxia associated with gluten sensitivity. *Brain*. 2001;124:1013–1019.
20. Paimela L, Kurki P, Leirisalo-Repo M, et al. Gliadin immune reactivity in patients with rheumatoid arthritis. *Clin Exp Rheumatol*. 1995;13:603–607.
21. Oberhuber G. Histopathology of celiac disease. *Biomed Pharmacother*. 2000;54:368–372.
22. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*. 1982;143:29–36.
23. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44:837–845.
24. Reeves GE, Squance ML, Duggan AE, et al. Diagnostic accuracy of coeliac serological tests: a prospective study. *Eur J Gastroenterol Hepatol*. 2006;18:493–501.
25. http://consensus.nih.gov/2004/2004_CeliacDisease118html.htm.
26. Dickey W, Hughes DF, McMillan SA. Reliance on serum endomysial antibody testing underestimates the true prevalence of coeliac disease by one fifth. *Scand J Gastroenterol*. 2000;35:181–183.
27. Abrams JA, Diamond B, Rotterdam H, et al. Seronegative celiac disease: increased prevalence with lesser degrees of villous atrophy. *Dig Dis Sci*. 2004;49:546–550.
28. Stern M. Comparative evaluation of serologic tests for celiac disease: a European initiative toward standardization. *J Pediatr Gastroenterol Nutr*. 2000;31:513–519.
29. Sugai E, Vazquez H, Nachman F, et al. Accuracy of testing for antibodies to synthetic gliadin-related peptides in celiac disease. *Clin Gastroenterol Hepatol*. 2006;4:1112–1117.
30. Prince HE. Evaluation of the INOVA diagnostics enzyme-linked immunosorbent assay kits for measuring serum immunoglobulin G (IgG) and IgA to deamidated gliadin peptides. *Clin Vaccine Immunol*. 2006;13:150–151.
31. Midhagen G, Aberg AK, Olcen P, et al. Antibody levels in adult patients with coeliac disease during gluten-free diet: a rapid initial decrease of clinical importance. *J Intern Med*. 2004;256:519–524.