

Diagnostic accuracy of coeliac serological tests: a prospective study

Glenn E.M. Reeves^a, Marline L. Squance^a, Anne E. Duggan^a, Rajathurai R. Murugasu^a, Robert J. Wilson^b, Richard C. Wong^b, Robert A. Gibson^a, Richard H. Steele^c, Wendy K. Pollock^d and the Multicentre Coeliac Study Group^{a-h}

Objective The best way to test serologically for coeliac disease (CD) remains controversial, with endomysial (EMA), transglutaminase (TTG), and gliadin antibodies (AGA) being assessed in various combinations with no apparent standardization. The objective of this study was to evaluate whether TTG-IgA +/- TTG-IgG could be used as a replacement for endomysial antibodies as a reliable screen for CD in patients presenting to a major Australian tertiary referral hospital for assessment of symptoms consistent with CD.

Methods Individuals referred for gastroscopic assessment of possible CD were prospectively evaluated by duodenal biopsy assessment. The following diagnostic methods were compared: dual-isotype transglutaminase (TTG-dual), combined-isotype transglutaminase (TTG-IgA + G), TTG-IgA, combined-isotype gliadin antibodies (AGA-IgA + G), AGA-IgA, and endomysial antibody assays. Clinical performance characteristics (sensitivity, specificity, area under the curve for receiver-operating characteristic analysis; AUROC) were assessed for all kits.

Results The correlation between transglutaminase kits was generally good, with the best transglutaminase kit demonstrating high correlation ($r=0.86$) with endomysial antibodies. A comparison of different types of endomysial antibody assays displayed variable diagnostic performance (sensitivity 61.90–68.42%; specificity 80.00–98.57%; AUROC 0.71–0.83). Sensitivity (90.48–92.31%), specificity (80.77–82.89%) and AUROC values (0.92–0.94) for dual-isotype transglutaminase kits displayed narrow ranges. AGA assays were less sensitive (AGA-IgA:

42.31–46.15%; AGA-IgG: 61.54%) and less specific (AGA-IgA: 85.09–87.73%; AGA-IgG: 82.46–84.09%). Dual-isotype transglutaminase testing was diagnostically equivalent to transglutaminase-IgA (AUROC 0.92 versus 0.91, $P=0.33$).

Conclusions Our study suggests that transglutaminase screening (using the IgA ± IgG isotype) is a sensitive and specific alternative to endomysial antibody testing in the serological assessment of CD. On the basis of our findings, AGA antibody testing no longer appears to be an essential part of the diagnostic strategy for adult CD. *Eur J Gastroenterol Hepatol* 18:493–501 © 2006 Lippincott Williams & Wilkins.

European Journal of Gastroenterology & Hepatology 2006, 18:493–501

Keywords: antibodies, coeliac, endomysial, intestinal biopsy, receiver-operating characteristic, sensitivity, serology, transglutaminase

^aHunter Area Pathology Service (HAPS), John Hunter Hospital, Newcastle, NSW, Australia, ^bRoyal Brisbane Hospital (RBH), Queensland, Australia, ^cCanterbury Health, Christchurch, New Zealand, ^dGribbles Pathology, Victoria, Australia, ^eInstitute of Medical and Veterinary Science (IMVS), South Australia, ^fQueensland Medical Laboratory, Queensland, Australia, ^gImmunology Department, Fremantle Hospital, Fremantle, Western Australia and ^hRoyal Prince Alfred Hospital, NSW, Australia

Correspondence and requests for reprints to Glenn E.M. Reeves, Hunter Area Pathology Service, John Hunter Hospital, Locked Bag 1, Hunter Region Mail Centre, NSW 2310, Australia
Tel: +61 2 49214029; e-mail: glenn.reeves@hnehealth.nsw.gov.au

Sponsorship: The study was partly supported by a grant from the Royal College of Pathologists of Australasia.

Received 14 October 2005 Accepted 14 January 2006

Introduction

Coeliac disease (CD) represents one of the best-understood examples of an environmentally triggered autoimmune disease, with genetic risk factors (especially HLA-DQ2/DQ8), environmental triggers (gliadin), and autoantigenic targets (tissue transglutaminase) being well defined [1]. Typically, CD targets the small intestine, and presents with malabsorption characterized by weight loss, abdominal distension, diarrhoea and steatorrhoea [2]. The clinical presentation of CD has

changed over recent decades, perhaps as a result of earlier detection with improved blood tests. Diarrhoea is a less frequent symptom than 10 years ago, and features such as unexplained iron deficiency and osteoporosis are more common [3]. This extension of the clinical spectrum of disease, in concert with improved diagnosis, has led to the revision of the prevalence of CD to approximately 1:100 [4]. The best way of testing for CD remains controversial, with endomysial (EMA), transglutaminase (TTG), and gliadin (AGA) antibodies being assessed

in various combinations with no apparent standardization [5,6].

The aim of the Multi-Centre Coeliac Serology Study was to define the optimal serological screening approach to CD in unselected patients referred for gastroscopic assessment of symptoms consistent with CD. Specific questions asked were: Can transglutaminase replace endomysial antibodies as a reliable screen for CD? Does combined or dual isotype transglutaminase testing (TTG-IgA + G or TTG-dual) of IgA-replete individuals enhance diagnostic sensitivity for CD compared with TTG-IgA testing alone? And does the addition of gliadin to the assessment of transglutaminase antibodies enhance the diagnostic sensitivity for CD compared with transglutaminase antibody testing alone?

Methods

The Multi-Centre Coeliac Serology Study was prospectively undertaken through seven major teaching and referral centres (Table 1) throughout Australia.

Sequential individuals referred to the central unit (Hunter) for gastroscopic assessment of possible CD between 25 March 2003 and 25 November 2004 were invited to participate in the study. At least one of a set of standardized clinical features consistent with CD (not including immunological markers, see Table 2 [7,8]) needed to be met. Study design, data collection and analysis and the reporting and interpretation of results were performed according to the STARD guidelines [9].

Patients and study design

Individuals aged between 4 and 90 years referred for endoscopic biopsy assessment of symptoms consistent with CD were eligible for the study. Three times a week, the endoscopy list of the John Hunter Hospital Day-Stay Gastroenterology Unit was reviewed for the presence of keywords consistent with a possible diagnosis of CD (Table 2). Individuals less than 4 years old or receiving anticoagulation treatment were excluded from the study.

Of the 2684 individuals presenting for gastroscopy during the study period, 1879 had been referred for endoscopic

Table 2 Clinical features consistent with coeliac disease [7,8]

Unexplained weight loss	Unexplained infertility
Short stature or 'failure-to-thrive' (in children)	Prolonged unexplained fatigue
Chronic diarrhoea	Family history of CD in first-degree relative(s)
Unexplained abdominal pain	Associated autoimmune conditions (type 1 diabetes, Sjogren's, lupus, ...)
Dyspeptic symptoms	Dermatitis herpetiformis-like rash
Alternating bowel habits	IgA deficiency
Constipation	Unexplained iron or folate deficiency
Non-degenerative arthralgias	Anaemia
Idiopathic epilepsy	Unexplained abnormal liver function tests
Idiopathic ataxia	
Idiopathic osteoporosis	

CD, Coeliac disease.

assessment of symptoms consistent with CD. Of these individuals, 894 were approached for participation in the study via a random selection process, and 732 agreed to be involved. We received 551 signed consent forms; 254 individuals were included in the final study (297 patients were excluded because of a failure to collect appropriate blood samples, no duodenal biopsy, or patient withdrawal). Of these individuals, 26 patients had biopsy-confirmed CD and 228 had no biopsy evidence of CD.

Patient consent was obtained to undergo the following procedures: endoscopy and biopsy (baseline), venepuncture for the collection of 20 ml blood (for the measurement of TTG-IgA, IgG TTG (TTG-IgG), dual-isotype TTG (TTG-dual), endomysial antibodies, total IgA, IgA anti-gliadin antibodies (AGA-IgA) and IgG anti-gliadin antibodies (AGA-IgG) at baseline.

This study was performed in accordance with the 1983 revision of the 1975 Helsinki declaration [10], and accorded with local guidelines for good clinical practice. All trial protocols, including any variations to the original study, were reviewed and approved by the Hunter Area Research Ethics Committee. Written informed consent was obtained from all patients.

Enzyme-linked immunosorbent assay and luminometry testing logistics

Enzyme-linked immunosorbent assay or luminometric testing for TTG-IgA and TTG-IgG as well as AGA-IgA and AGA-IgG was performed according to kit recommendations, as summarised in Table 3.

An expression of interest was circulated by mail and e-mail to all manufacturers offering kits for CD diagnosis. All methodologies available for detecting CD in Australasia at the time of study commencement were included. All participating kit manufacturers agreed to provide their kits free of charge to allow the study to proceed; however, the trial was designed and constructed independently by the authors, and there was no influence by the companies

Table 1 Testing centres participating in coeliac study

Study testing centre	Country
Hunter Area Pathology Service (HAPS), John Hunter Hospital, New South Wales	Australia
Immunology Fremantle Hospital (FHP), Western Australia	Australia
Royal Brisbane Hospital (RBH), Queensland	Australia
Queensland Medical Laboratory (QML), Queensland	Australia
Institute of Medical and Veterinary Science (IMVS), South Australia	Australia
Gribbles Pathology (GPM), Victoria	Australia
Royal Prince Alfred Hospital (RPAH), New South Wales	Australia
Canterbury Health (CHNZ), Christchurch	New Zealand

Table 3 Coeliac kit features

Kit and company	Antigen	Dilution	Incubation times	Calibrators	Conjugate	Substrate	Kit cut-off	ROC cut-off	Lab
Celicheck (Aeskulisa) Diagnostic Solutions	TTG-HRE	1:100	15;15;5	6	AH Ig (A & G) to HRP	TMB	15 U/ml	IgA: 6.19 IgG: 9.24 Dual: 783	HAPS
Aeskulisa Diagnostic Solutions	Alpha gliadin	1:100	30;15;15	6	AH Ig (A or G) HRP	TMB	15 U/ml	IgA: 12.1 IgG: 18.9	CHNZ
Bindzyme Immunotech	TTG-HRE	1:100	30;30;30	5	Rabbit HRP/AH Ig (A or G)	TMB	9 U/ml	3.0	FHP
Celkey ALS	TTG-HRE	1:101	30;30;10	6	AH Ig (A or G) to HRP	TMB	IgA: U/ml IgG: U/ml	6.0	IMVS
D-Tek Diagnostic technology	TTG-HRB	1:50	30;30;10	6	Rabbit HRP/AH Ig (A & G)	TMB	50 KU/l	IgA: 18.9 IgG: 13.6 Dual: 9.85	HAPS
Eurospiral Vital Diagnostics	TTG-HRE	1:25	60;60;30	1	Goat HRP/AH Ig (A or G)	TMB	7 KU/l	IgA: 2.1 IgG: 11.3	RBH
Euroimmun ESL Bioscience	TTG-HRB	1:200	30;30;15	3	Rabbit HRP/AH Ig (A or G)	TMB	20 KU/l	IgA: 3.4 IgG: 3.5	CHNZ
FIDIS TM Diagnostic Solutions	TTG-HRB	1:200	30;30	1	Goat/(PE)/AH IgA	N/a	20 AU/ml	13	GPM
FIDIS TM Diagnostic Solutions	Gluten (gliadin)	1:200	30;30	1	Goat/(PE)/AH IgA	N/a	20 AU/ml	12	GPM
Genesis Abacus	TTG-HRE	1:100	30;30;30	IgA-5 IgG-4	Rabbit HRP/AH Ig (A or G)	TMB	7 KU/l	IgA: 6 IgG: 10	QML
Inova Integrated Sciences	TTG-HEP	1:100	30;30;30	1	Goat HRP/AH Ig (A or G)	TMB	20 U	IgA: 10.8 IgG: 8.1	HAPS
Inova Integrated Sciences	Gliadin	1:100	30;30;30	1	Goat HRP/AH Ig (A or G)	TMB	20 U	IgA: 15 IgG: 23	GPM
Medical Innovations Vital Diagnostics	Gliadin	1:100	15;15;15	1	HRP (AH) IgA and IgG	ABTS	Child IgA <25 IgG <46 Adult IgA <34 IgG <42	IgA: 25 IgG: 36	QML
Orgentec Immunodiagnosics	TTG-HRM	1:100	30;15;15	6	Rabbit HRP/AH Ig (A or G)	TMB	10 KU/l	IgA: 2.78 IgG: 9.5	IMVS
Endomysial antibodies Immco	PSM	1:10 +	N/a	-	FITC goat AH Ig (polyvalent)	N/a	Negative	≥ 1:40	RPAH

AH, Anti-human; FIDISTM coeliac IgA panel; FITC, fluorescein isothiocyanate; HEP, human erythrocyte (purified); HRP, human recombinant (baculovirus); HRE, human recombinant (*E. coli*); HRM, human recombinant (myeloma kidney cell line); HRP, horseradish peroxidase; PE, phycoerythrin; PSM, primate smooth muscle; ROC, receiver-operating characteristic; TTG, transglutaminase.

All kits used 3,3',5,5'-tetramethylbenzidine substrate.

ROC-derived cut-off determined using value providing >95% sensitivity.

Fid is a luminometric (laser-addressable-particle) technology: only IgA-isotype testing available for TGA and gliadin (AGA) on this methodology.

upon the study process, analysis or publication. All test analyses occurred in batches of approximately 80 samples, with each batch being sent to a central centre for aliquoting (RBH) to ensure uniform sample characteristics and volumes, then forwarded to each participating laboratory. The eight teaching hospital-based laboratories participating in sample analysis were chosen to perform a given test based upon their experience using the test method in question.

The data were retained and analysed by the study coordinators in a confidential password-protected database. Table 3 summarizes laboratories, testing kits and their suppliers, along with details of the kit methods.

Endomysial antibody subanalysis

Endomysial antibodies were assessed by indirect immunofluorescence using 'primate smooth muscle' substrate (Immco, Buffalo, USA), with results of 1:10 or greater being reported as positive. A single blinded observer with experience in coeliac immunofluorescence read all endomysial slides. The laboratory assessing endomysial antibodies performed dual-isotype endomysial antibody, and endomysial antibody samples were tested for masking as a result of anti-smooth muscle antibodies [11].

Because the term 'primate smooth muscle' has been used generically in some publications to refer to both monkey distal oesophagus and monkey bladder substrates [12], it was decided that further analysis of a random subset of 161 samples from the original sample set would be assessed in parallel to compare IgA-endomysial antibodies on primate distal oesophagus (EMA-ME-IgA), IgA-endomysial antibodies on monkey bladder (EMA-MB-IgA), and dual-endomysial antibodies on monkey bladder (dual-EMA-MB). The trial coordinators regarded this as an important enhancement of this study, given the fact that the diagnostic performance of each of these 'endomysial antibodies' may be different. EMA-ME-IgA and EMA-MB-IgA were analysed using an indirect immunofluorescent method (Euroimmun, Luebeck, Germany).

Duodenal biopsy analysis

Three samples were taken from the second part of the duodenum during the endoscopic assessment of patients with possible CD. Two pathologists performed histological assessment in a blinded fashion using modified Marsh criteria.

Statistics

The required sample size of 23 patients with CD and 217 individuals without CD was calculated to establish a difference in the area under curve for receiver-operating characteristic analysis (AUROC) of 0.05 (from 0.9 to

0.95) with 95% confidence and 80% power, assuming a ratio of non-CD:CD of 9:1.

Statistical analyses were performed using Stata version 8 (Stata Corporation, College Station, Texas, USA) and JMP Version 5.1 (SAS, Cary, North Carolina, USA). Clinical and measured variables were summarized as means $\pm$ SD (with range) for continuous data, and as proportions (with percentages) for categorical information. Differences between these summary statistics were analysed using the Student *t* and Mann–Whitney–Wilcoxon tests for continuous variables and the χ^2 test for categorical variables. The Fisher exact test was employed when the minimum expected value with χ^2 testing was less than 5. Linear regression analysis was used to correlate measured parameters. Nominal logistic regression methods were used to calculate the combined AUROC performance of different assays used in combination [13]. A probability value of $P < 0.05$ (two-sided) was considered significant. AUROC values for different kits were compared using standard methodology, which corrected for possible correlations between areas. Kit comparisons were undertaken using Pearson's correlation, Bland–Altman analysis and Deming regression statistics [14].

Diagnostic assay characteristics

Sensitivity was calculated as the number of true positive results divided by the total number of individuals with CD, with specificity equal to the proportion of true negatives within the non-CD group. A positive likelihood ratio (PLR) was calculated using the formula: sensitivity $\div$ (1 – specificity), with these values derived at the critical cut-off when the optimal diagnostic sensitivity and specificity were achieved. Similarly, a negative

likelihood ratio was calculated as follows: (1 – sensitivity) $\div$ specificity.

Results

Patient demographics

A total of 551 patients were recruited into the study between March 2003 and November 2004. Of these individuals, complete data (baseline endoscopic biopsy results, TTG-IgA, TTG-IgG, total IgA, endomysial antibodies (dual-isotype), AGA-IgA, AGA-IgG) were available for 254. The basic demographics on study participants appear in Table 4.

Assay characteristics

Kit correlation results are shown in Tables 5, 6 and 7.

Diagnostic test characteristics

Diagnostic characteristics for each of the available enzyme-linked immunosorbent and luminometry-based assays are listed in Table 8. A positive result on a kit with a PLR of more than 5 indicates that the likelihood of CD is over 20%

Table 4 Features of study participants

Age [mean $\pm$ SD (range)]	47.8 $\pm$ 17.7 (6–82)
Paediatric (age 4–12) (paediatric %)	9 (3.5%)
Sex (female: male) (male %)	177:77 (30.3%)
Non-coeliacs:coeliacs (coeliac %)	228:26 (10.24%)
IgA-deficient (%) – non-coeliacs:coeliacs (%)	27 (10.6%) 24:3 (11.5%)

Table 5 Dual TTG (IgA + IgG) correlations

	Celicheck	D-Tek	Endomysial antibody
Celicheck	1.00		
D-Tek	0.92	1.00	
Endomysial antibody	0.81	0.89	1.00

Table 6 TTG-IgA correlations

	Aeskulisa	Inova	Orgentec	Bindazyme	Eurospital	Celikey	Genesis	Euroimmune	Fidis	D-Tek
Aeskulisa	1.00									
Inova	0.87	1.00								
Orgentec	0.82	0.87	1.00							
Bindazyme	0.83	0.75	0.82	1.00						
Eurospital	0.87	0.90	0.80	0.72	1.00					
Celikey	0.83	0.84	0.97	0.81	0.81	1.00				
Genesis	0.79	0.73	0.74	0.88	0.74	0.75	1.00			
Euroimmun	0.84	0.79	0.89	0.86	0.75	0.86	0.77	1.00		
Fidis	0.84	0.81	0.74	0.76	0.80	0.78	0.68	0.75	1.00	
D-Tek	0.64	0.33	0.18	0.46	0.43	0.24	–0.02	0.40	0.54	1.00

Table 7 AGA kit correlations

	AGA-IgA				AGA-IgG		
	Inova	Medical Innovations	Aeskulisa	Fidis	Inova	Medical Innovations	Aeskulisa
Inova	1.00				1.00		
Medical Innovations	0.62	1.00			0.72	1.00	
Aeskulisa	0.90	0.74	1.00		0.55	0.90	1.00
Fidis	0.37	0.48	0.42	1.00			

Table 8 Diagnostic test characteristics

Kit	CD: % positive (sensitivity)	No CD: % negative (specificity)	AUROC (95% CI)	PLR	NLR
TTG kits					
Celicheck dual	92.31	82.89	0.92 (0.86–0.99)	5.40	0.09
D-Tek dual	90.48	80.77	0.94 (0.87–1.00)	4.70	0.12
Aeskulisa IgA	84.62	86.40	0.89 (0.80–0.98)	6.22	0.18
Aeskulisa IgG	84.62	89.04	0.88 (0.78–0.97)	7.72	0.17
Bindazyme IgA	80.77	93.69	0.91 (0.84–1.00)	12.80	0.21
Bindazyme IgG	65.38	85.59	0.79 (0.67–0.91)	4.54	0.40
Celikey IgA	80.77	93.86	0.87 (0.76–0.98)	13.15	0.20
Celikey IgG	66.67	82.05	0.80 (0.70–0.91)	3.71	0.41
D-Tek IgA	16.67	97.08	0.62 (0.52–0.71)	5.71	0.86
D-Tek IgG	16.67	89.78	0.57 (0.48–0.67)	1.63	0.93
Euroimmun IgA	80.77	83.70	0.85 (0.75–0.96)	4.96	0.23
Euroimmun IgG	76.92	85.09	0.87 (0.80–0.95)	5.16	0.27
Eurospital IgA	92.31	81.14	0.92 (0.83–1.00)	4.89	0.09
Eurospital IgG	76.92	85.53	0.89 (0.81–0.96)	5.31	0.27
Fidis IgA	88.46	84.65	0.89 (0.81–0.98)	5.76	0.14
Genesis IgA	73.08	94.74	0.85 (0.76–0.95)	13.88	0.28
Genesis IgG	69.23	82.89	0.78 (0.67–0.90)	4.05	0.37
Inova IgA	76.92	87.22	0.88 (0.79–0.97)	6.02	0.26
Inova IgG	42.31	85.96	0.70 (0.60–0.82)	3.01	0.67
Orgentec IgA	88.46	83.77	0.91 (0.82–1.00)	5.45	0.14
Orgentec IgG	69.23	81.5	0.85 (0.76–0.93)	3.74	0.38
AGA kits					
Aeskulisa IgA	42.31	87.73	0.71 (0.59–0.82)	3.45	0.66
Aeskulisa IgG	61.54	84.09	0.82 (0.74–0.89)	3.87	0.46
Fidis IgA	46.15	85.09	0.69 (0.57–0.81)	3.10	0.63
Inova IgA	46.15	85.09	0.75 (0.65–0.86)	3.10	0.63
Inova IgG	61.54	82.46	0.79 (0.70–0.88)	3.51	0.47
Medical Innovations IgA	42.31	85.53	0.66 (0.54–0.78)	2.92	0.67
Medical Innovations IgG	61.54	82.46	0.80 (0.73–0.89)	3.51	0.47

AGA, Gliadin; AUROC, area under the curve for receiver-operating characteristic; CD, coeliac disease; CI, confidence interval; NLR, negative likelihood ratio; PLR, positive likelihood ratio; TTG, transglutaminase.

Results for sensitivity, specificity, PLR and NLR were calculated using receiver-operating characteristic-adjusted cut-offs for all kits.

for an individual in whom the pre-test probability of CD is only 5%. The best three kits obtained PLR values greater than 5 and AUROC values greater than 0.90.

IgA-deficient individuals

Twenty-seven individuals (10.6%) had partial or total IgA deficiency. IgA reference intervals were age specific, according to published data [15]. Three of these individuals (11.5%) were diagnosed with CD (total IgA levels < 0.07, 0.51, 0.73). The total IgA level was also analysed using standard logistic regression analysis to assess whether the measurement of total IgA levels added any diagnostic information. Total IgA levels did not add diagnostic information in these analyses. In IgA-deficient individuals, the performance of TTG-IgG and AGA-IgG testing was assessed to evaluate the role of IgG-based serology in this setting. TGG-IgG increased diagnostic sensitivity for CD in IgA-deficient individuals (sensitivity of 100% for dual-isotype transglutaminase versus 66.67% for TTG-IgA). However, despite being IgA-deficient, a significant proportion of CD subjects (67%) were able to mount TTG-IgA responses, although this was only seen in the two patients displaying partial rather than absolute IgA deficiency.

Paediatric populations

Nine (3.5%) individuals were between 4 and 12 years of age (inclusive). Of these patients two were IgA deficient. There was no significant difference in the analytical

method performance in this subgroup compared with the total cohort, although the total numbers were small.

Endomysial antibody subanalysis

A subset of 140 individuals without CD plus 21 of the patients with known CD was analysed. Results (Table 9) indicated that dual-endomysial antibodies on monkey bladder (dual-EMA-MB) performed the best of the endomysial antibody methods tested in this study.

Comparison of generic approaches

Generic approaches to diagnosis were analysed. Table 10 displays the diagnostic characteristics for each approach, using the results from the best performing kit. Indices of correlation between the severity of duodenal histological change and serological results were calculated using Spearman's method. Correlation coefficients for each of these antibodies reflected a low association between serology and histology for TTG-dual (0.45), TTG-IgA (0.41), and endomysial antibodies (0.43).

Dual-isotype and combined-isotype transglutaminase testing

Sensitivity, specificity and AUROC values for the best dual isotype transglutaminase kit were 92.31%, 82.89% and 0.92, respectively, compared with values of 88.46%, 83.77% and 0.91, respectively, for the best TTG-IgA

Table 9 Endomysial method analysis

Kit	CD: % positive (sensitivity)	No CD: % negative (specificity)	AUROC (95% CI)	PLR	NLR
Dual-EMA-MB	68.42	98.57	0.83 (0.73–0.94)	47.89	0.32
EMA-ME-IgA	61.90	92.14	0.77 (0.66–0.88)	7.88	0.48
EMA-MB-IgA	61.90	80.00	0.71 (0.60–0.82)	3.10	0.67

AUROC, Area under the curve for receiver-operating characteristic; CD, coeliac disease; CI, confidence interval; Dual-EMA-MB, dual-endomysial antibodies on monkey bladder; EMA-ME-IgA, IgA-endomysial antibodies on primate distal oesophagus; EMA-MB-IgA, IgA-endomysial antibodies on monkey bladder; NLR, negative likelihood ratio; PLR, positive likelihood ratio.

Table 10 Diagnostic characteristics of kits assessed

Kit	CD: % positive (sensitivity)	No CD: % negative (specificity)	AUROC (95% CI)	PLR	NLR
EMA	68.42	98.57	0.83 (0.73–0.94)	47.89	0.32
TTG-IgA	88.46	83.77	0.91 (0.82–1.00)	5.45	0.14
TTG-IgA (Lum)	88.46	84.65	0.89 (0.81–0.98)	5.76	0.14
TTG-IgG	84.62	89.04	0.88 (0.78–0.97)	7.72	0.17
Dual-TTG	92.31	82.89	0.92 (0.86–0.99)	5.40	0.09
AGA-IgA	46.15	85.09	0.75 (0.65–0.86)	3.10	0.63
AGA-IgA (Lum)	46.15	85.09	0.69 (0.57–0.81)	3.10	0.63
AGA-IgG	61.54	84.09	0.82 (0.74–0.89)	3.87	0.46
IgA-Def	15.79	89.47	0.53	1.50	0.94

AGA, Gliadin; AUROC, area under the curve for receiver-operating characteristic; CD, coeliac disease; CI, confidence interval; EMA, endomysial antibody; NLR, negative likelihood ratio; PLR, positive likelihood ratio; TTG, transglutaminase.

Because of its unique characteristics, EMA has been analysed separately from other dual isotype methods. Results incorporated values for the best kit in each grouping. D-Tek IgA was dropped because of aberrant results.

method (no significant difference, $P = 0.33$). The combined performance of test methods offering TTG-IgA and TTG-IgG separately was also assessed using nominal logistic regression methods. Calculated combined AUROC performances (TTG-IgA plus TTG-IgG) were as follows: Aeskulisa 0.93; Inova 0.90; Orgentec 0.92; Bindazyme 0.87; Eurospital 0.95; Celikey 0.91; Genesis 0.90; Euroimmun 0.91; and D-Tek 0.56.

Discussion

It is increasingly accepted that the measurement of IgA antibodies to human tissue transglutaminase (TTG-IgA) offers equal or superior performance to endomysial antibody (EMA-IgA) assessment using monkey substrates [16,17]. This is not surprising, given the fact that the dominant target antigen of endomysial antibodies is transglutaminase, and both methods measure 'transglutaminase antibodies' [18]. In our study, TTG-IgA screening was superior to endomysial antibody testing. Some authors argue that endomysial antibody still plays a role in view of its slightly greater diagnostic specificity in some studies [19], and this has led to the use of a 'two-tier' serological screening system in some centres, with transglutaminase followed by endomysial antibodies [20]. Our study does not support this approach, which will only add to the expense and diagnostic delay, with diagnostic sensitivity reduced by more than 20%.

As witnessed by the imperfect sensitivity of even the best transglutaminase kits, serological studies are not a substitute for tissue biopsy assessment when CD is

strongly suspected. The prospective nature of this study may partly explain why the sensitivity of serology for diagnosing CD was only approximately 90% compared with previous retrospective analyses suggesting figures of 95–99%.

Picarella *et al.* [21] have identified an IgA-replete group of individuals for whom CD is only diagnosed serologically using IgG endomysial antibodies (EMA-IgG), and this phenomenon has been replicated for IgG antibodies to transglutaminase (TTG-IgG) [22]. This finding, in addition to the recognition of a heightened prevalence of IgA deficiency (IgA-D) in CD, has led to studies exploring the incremental diagnostic utility of dual-isotype screens (TTG-IgA plus TTG-IgG or dual EMA) for CD.

Dual-isotype and combined-isotype testing of TTG-IgA plus TTG-IgG offered equivalent diagnostic reliability for CD in our study. In comparison with AGA-IgG, TGG-IgG displayed acceptable sensitivity and specificity. For the 10.6% of individuals with reduced IgA levels, the dual or combined isotype-TGA testing approach was superior to either isotype measured in isolation.

Whereas some studies suggest that IgA and IgG antibodies to gliadin (AGA-IgA, AGA-IgG) still play a role in CD diagnosis, increasing evidence suggests that AGA-IgG should be discarded in view of its poor specificity. AGA adds little to transglutaminase-based methods in terms of diagnosing CD, although some

studies suggest that gliadin antibody assessment might add to diagnostic sensitivity for extra-intestinal forms of gluten sensitivity, especially neurological syndromes [23]. The addition of AGA cannot be recommended on the basis of our study results. None of our study participants suffered from neurological manifestations as a primary presentation of their gluten sensitivity; therefore, we cannot comment on any role for AGA assessment in this setting. However, given the rarity of such syndromes, our study suggests that AGA testing can largely be abandoned for routine CD testing.

The equivalence between endomysial antibodies and transglutaminase is not complete, with discordances between the two serological methods having been attributed to a range of antigenic targets of uncertain importance, including calreticulin [24] and F-actin. The measurement of IgA antibodies to actin is of particular interest currently, as some early reports suggest that titres of this antibody, reportedly specific for CD, might be unique among CD serologies in mirroring duodenal histological changes [25]. This particular question is being assessed through a subanalysis of the current study. This study was not designed to assess the importance of non-transglutaminase antigens in CD; however, it can be deduced from our results that the measurement of antibodies to such antigens does not add significantly to the diagnostic yield for CD. This appears particularly true given the superiority of transglutaminase methods over the endomysial antibody assay.

This study compared a number of methods for analysing endomysial antibodies. Previous articles suggest that dual or combined isotype testing may penalize endomysial antibody performance through the creation of false-positive results by the inclusion of EMA-IgG [21]. This was not shown to be the case in our study, with the dual-EMA-MB approach offering high specificity for CD (98.57%), albeit at the expense of a loss of sensitivity. In addition, the demonstration of the superior diagnostic performance of transglutaminase compared with endomysial antibodies is not unique to our study, and adds to a growing body of evidence favouring transglutaminase as a diagnostic tool in CD.

Not all 'endomysial antibodies' are the same. Endomysial structures are detectable in a range of tissue substrates, including monkey bladder and human umbilical cord. It is important to ensure that kits labelled 'primate smooth muscle' actually use monkey distal oesophagus, as the diagnostic performance of endomysial antibodies has not been extensively validated on other tissues, such as monkey bladder. Our study demonstrated the substrate-dependent diagnostic performance of endomysial antibody serology with sensitivity ranging from 61.90 to 68.42% and specificity ranging from 80.00 to 98.57%. For

this reason we recommend caution in choosing endomysial antibody tests.

As a measure of diagnostic test accuracy, the AUROC parameter is highly useful, with values greater than 0.90 indicating a superior test performance. Cut-offs should be validated for each laboratory performing under local conditions on local populations, assessing patients with CD, disease controls (especially inflammatory bowel disease and systemic autoimmune conditions) and normal individuals. The optimization of assay cut-offs using receiver-operating characteristic analysis improves test performance [26]. However, two tests may display significantly different practical usefulness while possessing identical AUROC. For example, in our study, two kits possessed good AUROC values (0.91), but one kit would be preferred for routine diagnosis because of its greater specificity at a cut-off level providing high sensitivity for detecting CD. In other words, two tests with identical AUROC values may demonstrate different PLR values for a given sensitivity cut-off. It was found that many kits performed poorly, with 12 kits displaying AUROC values of less than 0.80. In our opinion, measures of diagnostic test performance combining such features as AUROC, sensitivity and PLR provide a more comprehensive and balanced picture of overall test reliability, and should be utilized in all studies comparing different diagnostic methodologies.

A positive result on a kit with a PLR of over 5 indicates that the likelihood of CD is more than 20% for an individual with a pre-test probability as low as 5%, a reasonable threshold to justify the inconvenience of endoscopic assessment and biopsy. This represents a potentially powerful approach to detecting CD at a population level, a practice that is arguably justified by the conformance of transglutaminase testing with World Health Organization criteria for an 'ideal' screening test [27]. Given the residual potential for high numbers of falsely positive results, population screening for CD is not currently recommended. A 'case-finding' approach based on the presence of features or syndromes associated with CD is justified by the availability of reliable serological methods [28].

Our study design reduced ascertainment bias, as the inclusion criteria were broad, clinically based, and not overtly dependent upon previous serological results. The breadth of the criteria used for inclusion in the study reflects the fact that CD symptoms are often non-specific with a poor predictive value. It could be argued that individuals who submit to a study protocol involving endoscopy and biopsy are more likely to have significant diseases such as CD. This risk was minimized by the fact that our institution exercises a low threshold for biopsy assessment of possible CD, and will explore this possibility in a wide variety of clinical settings. The

generally lower sensitivities of all kit methods for diagnosing CD in our study compared with previous publications supports this contention [29]. However, some degree of selection bias is likely to exist in this study. There was a high drop-out rate, related largely to the reluctance to undertake serial venepuncture, which may artefactually bias the patient mix towards more severe disease and a greater likelihood of CD. Expectation bias may have existed because pathologists performing blinded analysis were aware that testing was occurring in the context of a coeliac diagnostic study. Individuals referred for biopsy assessment of CD are likely to have had previous positive CD serology, introducing referral bias. Nothing short of a general practice-based study comparing 'up-front' biopsy with initial serology followed by biopsy could attempt to address this question, and these resources were not available to our study centres. Given the small proportion of paediatric patients in our study, the generalization of these recommendations outside of adult populations should be undertaken with caution.

All the serological methods assessed in this study were judged against the rigid criterion of biopsy-demonstrated duodenal histological abnormality, enhancing the uniformity of the study. Despite our inclusion of subtle histological changes, this limitation potentially excluded a significant number of individuals with 'latent' CD, dermatitis herpetiformis [30], gluten-associated steatohepatitis [31], and other conditions [32]. In most of the studies of CD serology to date, positive antibody results in the absence of demonstrable bowel pathology have been labelled falsely positive, thus reducing the diagnostic specificity of serology. It is possible that many of these patients with 'normal' small bowel biopsies but high levels of endomysial or transglutaminase antibodies could be included in the category of latent or overt CD if other relevant clinical features are evident (e.g. unexplained iron deficiency or transaminitis), particularly if these occur in the setting of an appropriate family history or genotype, IgA deficiency, or other autoimmune conditions such as type 1 diabetes or autoimmune thyroiditis [33]. The potential causes of discrepancy between small bowel histology and CD serology are well known, and include premature dietary manipulation and the use of concurrent immunomodulation therapy (including prednisone) [10].

There is growing enthusiasm to include HLA-genotyping analysis in the diagnostic work-up of suspected CD. Although our study did not include genetic testing, the positive likelihood values displayed by our study methods indicate that genetic testing would have added little in terms of diagnostic sensitivity, and current evidence only indicates a clear role for such genetic testing in the clarification of ambiguous serohistological results, or in paediatric populations less than 2 years of age in which

serology is known to be less reliable [34]. Given the high negative predictive value of testing for the DQ2/DQ8 allele [35], genetic testing may also be useful in clarifying the status of patients who cannot undertake gluten re-challenge, or patients receiving concurrent immunomodulation.

The presence of IgA deficiency has been reported to increase the probability of CD 10-fold [7,8], in addition to creating the potential for falsely negative IgA-based CD serological test results. Our data indicate, however, that total IgA measurement does not provide enhanced diagnostic sensitivity for CD when dual-isotype transglutaminase testing is employed. Of interest is the fact that a large proportion of individuals with IgA deficiency (67%) demonstrated significant elevation of TTG-IgA levels in the presence of CD, although this was not witnessed in the single individual with absolute IgA deficiency (<0.07). This implies that some residual capacity for antigen-specific IgA production may persist in the presence of partial, but not complete, IgA deficiency.

In conclusion, our study suggests that dual or combined-isotype transglutaminase screening can be advocated as a sensitive and specific alternative to endomysial antibody testing in the serological assessment of CD. On the basis of our findings, AGA antibody testing no longer appears to be an essential part of the diagnostic strategy for adult CD.

Acknowledgements

The authors would like to express their gratitude to the following kit manufacturers who allowed this study to occur, through the provision of free diagnostic kits: Vital Diagnostics, Immunotech, Immunodiagnostics, Australian Laboratory Sciences, Diagnostic Solutions, Diagnostic Technology, ESL Bioscience, Immco, Abacus, and Integrated Sciences. They are also indebted to their other Hunter and Australasian colleagues for their collaboration and commitment, in particular Robert Wilson for his exhaustive work in preparing aliquots for generalized analysis. Special thanks also go to the other member of the Multicentre Coeliac Group: K. Lemmert, M. Hendle, M. Trevisin, P. Bhathal, J. Cassidy, D. Willis, A. Williams, S. Adelstein, R. Garsia, L. Weinholt, E. Pennington, P. Davidson, G. Roy, D. Gillis, D. Mallon, J. Attia, L. Irwig, K. Baumgart, S. Shepherd, W. Dimech, M. Cherry, D. Edmunds, R. Eastlake, S. Ong, P. Crock, M. Spellerberg, C. Clarke, S. Hall, A. Hall, K. Kirkwood, K. Brown and T. Taylor. The centres involved in the study included John Hunter Hospital, Newcastle (JHH) (Primary Research Centre), Royal Brisbane Hospital (RBH) (Aliquoting Centre), Queensland Medical Laboratory, Institute of Medical and Veterinary Science, Gribbles Pathology, Royal Prince Alfred Hospital, and Canterbury Health. The study team would also like to

express their gratitude to the gastroenterologists who kindly agreed to participate in the Multicentre Coeliac Study, of special note: Drs Anne Duggan, Robert Gibson, Sia Ong, Gerard Roy, Dakshesh Vakil, Douglas Routley, Neil Porter, Colin Thursby, Patricia Davidson and Rajendra Kumar of John Hunter Hospital, NSW, Australia.

Conflict of interest

None of the members of the Multicentre Coeliac Study Group has any commercial bond with the manufacturers of the kits tested. Although individual kits were identifiable by the testing centres performing the assays, all antibody result determinations were centrally de-identified, collated and analysed by the study coordinators, Dr Glenn Reeves and Ms Marlene Squance, resulting in the testing centres participating being blinded to individual kit results. Study participants' medical details and sample identification were protected via a unique coded database in accordance with Commonwealth privacy laws and the NSW Health Records and Information Privacy Act 2002 of Australia.

Authors' contributions

This article is based upon the work of the Multicentre Coeliac Study Group over the time period of November 2003 to November 2004. Patient recruitment and sample collection occurred at John Hunter Hospital, Newcastle, NSW, Australia, and was coordinated by Dr Glenn Reeves, Ms Marlene Squance, Dr Anne Duggan and Dr Robert Gibson. Serological kits utilized within the study were provided free of charge by the Australian and New Zealand manufacturers of these kits.

References

- Farrell RJ, Kelly CP. Current concepts: celiac sprue [Review article]. *N Engl J Med* 2002; **346**:180–188.
- Barr GD, Grehan MJ. MJA practice essentials gastroenterology: coeliac disease. *Med J Aust* 1998; **169**:109–114.
- Fasano A, Catassi C. Current approaches to diagnosis and treatment of celiac disease: an evolving spectrum. *Gastroenterology* 2001; **120**: 636–651.
- Schuppan D, Hahn EG. IgA anti-tissue transglutaminase: setting the stage for coeliac disease screening. *Gastroenterol Hepatol* 2001; **13**:635–637.
- Wong RC, Steele RH, Reeves GEM, Wilson RJ, Pink A, Adelstein S. Antibody and genetic testing in celiac disease. *Pathology* 2003; **35**: 285–304.
- Stern M. Comparative evaluation of serological tests for celiac disease: a European initiative toward standardization. *J Pediatr Gastroenterol Nutr* 2000; **31**:513–519.
- Cataldo F, Marino V, Ventura A, Bottaro G, Corazza GR. Prevalence and clinical features of selective immunoglobulin A deficiency in coeliac disease: an Italian multicentre study. Italian Society of Paediatric Gastroenterology and Hepatology (SIGEP) and 'Club del Tenue' Working Groups on Coeliac Disease. *Gut* 1998; **42**:362–365.
- Reeves GEM. Coeliac disease: against the grain. *Intern Med J* 2004; **34**:522–525.
- Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, *et al.* Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Clin Chem* 2003; **49**:1–6.
- The World Medical Association Declaration of Helsinki. Ethical principles for medical research involving human subjects. Available at: <http://www.wma.net/e/policy/b3.html>. Accessed: 25 January 2003.
- Lasagni D, Ferrari R, Lapini M. Unmasking anti-endomysial antibodies in coeliac subjects positive for anti-smooth muscle antibodies. *Acta Paediatr* 1999; **88**:462–464.
- Kumar V, Jarzabek-Chorzelska M, Sulej J, Karnewska K, Farrell T, Jablonska S. Celiac disease and immunoglobulin A deficiency: how effective are the serological methods of diagnosis? *Clin Diag Lab Immunol* 2002; **9**:1295–1300.
- Hosmer DW, Lemeshow S. *Applied logistic regression*. New York: John Wiley and Sons; 1989.
- Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Meth Med Res* 1999; **8**:135–160.
- HAPS Handbook. Normal reference ranges for serum immunoglobulin A. June 2005. Available at: www.haps.nsw.gov.au/education/handbook. Accessed: 18 June 2005.
- Reeves GEM, Burns C, Hall ST, Gleeson M, Lemmert K, Clancy RL. The measurement of IgA and IgG transglutaminase antibodies in celiac disease: a comparison with current diagnostic methods. *Pathology* 2000; **32**: 181–185.
- Wolters V, Vooijs-Moulaert AF, Burger H, Broomans R, De Schryver J, Rijkers G, Houwen R. Human tissue transglutaminase enzyme linked immunosorbent assay outperforms both the guinea pig based tissue transglutaminase assay and anti-endomysium antibodies when screening for celiac disease. *Eur J Paediatr* 2002; **161**:284–287.
- Dieterich W, Ehnis T, Bauer M, Donner P, Volta U, Riecken EO, *et al.* Identification of tissue transglutaminase as the autoantigen of celiac disease. *Nat Med* 1997; **3**:797–801.
- Bazzigalupi E, Lampasona V, Barere G, Venerando A, Bianchi C, Chiumello G, *et al.* Comparison of tissue transglutaminase-specific antibody with established antibody measurements for celiac disease. *J Autoimmun* 1999; **12**:51–56.
- Horvath K, Hill ID. Anti-tissue transglutaminase antibody as the first line screening for celiac disease: good-bye antigliadin tests? *Am J Gastroenterol* 2002; **97**:2702–2704.
- Picarelli A, Tola MD, Sabbatella L, Mastracchio A, Trecca A, Gabrielli F, *et al.* Identification of a new celiac disease subgroup: antiendomysial and anti-transglutaminase antibodies of IgG class in the absence of selective IgA deficiency. *J Intern Med* 2001; **249**:181–188.
- Cataldo F, Lio D, Marino V, Picarelli A, Ventura A, Corazza GR, and the Working Groups on Celiac Disease of SIGEP and Club del Tenue. IgG (1) antiendomysium and IgG antitissue transglutaminase (anti-TTG) antibodies in celiac patients with selective IgA deficiency. *Gut* 2000; **47**:366–369.
- Srikanth VK, Anderson JFI, Donnan GA, Saling MM, Macdonnell RAL. Multiple sclerosis and occult gluten sensitivity. *Neurology* 2005; **64**: 933–934.
- Sanchez D, Tuckova L, Sebo P, Michalak M, Whelan A, Sterzl I, *et al.* Occurrence of IgA and IgG autoantibodies to calreticulin in coeliac disease and various autoimmune diseases. *J Autoimmun* 2000; **15**:441–449.
- Carroccio A, Brusca I, Iacona G, Prima LD, Teresi S, Pirrone G, *et al.* Anti-actin antibodies in celiac disease: correlation with intestinal mucosa damage and comparison of ELISA with the immunofluorescence assay. *Clin Chem* 2005; **51**:917–920.
- Hilden J. What properties should an overall measure of test performance possess? *Clin Chem* 2005; **51**:471–472.
- Green PHR, Jabri B. Coeliac disease. *Lancet* 2003; **362**:383–391.
- Young EH, Wareham NJ. Screening for coeliac disease: what evidence is required before population programmes could be considered? *Arch Dis Child* 2004; **89**:499–500.
- Wong RC, Wilson RJ, Steele RH, Radford-Smith G, Adelstein S. A comparison of thirteen guinea pig and human anti-tissue transglutaminase antibody ELISA kits. *J Clin Pathol* 2002; **55**:488–494.
- Dieterich W, Laag E, Bruckner-Tuderman L, Reunala T, Karpati S, Zagoni T, *et al.* Antibodies to tissue-transglutaminase as serologic markers in patients with dermatitis herpetiformis. *J Invest Dermatol* 1999; **113**:133–136.
- Duggan JM. Coeliac disease: the great imitator. *Med J Aust* 2004; **180**:524–526.
- Reeves GEM. Coeliac disease [Letter]. *Int Med J* 2005; **35**:375.
- Book LS. Diagnosing coeliac disease in 2002: who, why and how? *Pediatrics* 2002; **109**:952–954.
- North American Society for Pediatric Gastroenterology, Hepatology and Nutrition. Guideline for the diagnosis and treatment of celiac disease in children: recommendations of the North American Society for Pediatric Gastroenterology, Hepatology and Nutrition. *J Pediatr Gastroenterol Nutr* 2005; **40**:1–19.
- Kaukinen K, Partanen J, Maki M, Collin P. HLA-DQ typing in the diagnosis of celiac disease. *Am J Gastroenterol* 2002; **97**:695–699.