Bibliographic reference	Burgin-Wolff (2013): Intestinal biopsy is not always required to diagnose coeliac disease: a retrospective analysis of combined antibody tests
Study type	Cohort (retrospective)
Study quality	1. Could the selection of patients have introduced bias? NO – consecutive patients recruited 2. Is there concern that the included patients do not match the review question? NO – all patients suspected of CD 3. Could the conduct or interpretation of the index test have introduced bias? NO – N/A 4. Is there concern that the index test, its conduct, or interpretation differ from the review question? NO – index test as specified in protocol 5. Could the reference standard, its conduct, or its interpretation have introduced bias? NO 6. Is there concern that the target condition as defined by the reference standard does not match the review question? NO – target condition matches review question 7. Could the patient flow have introduced bias? YES – 'For an undefined period of time patients were sometimes selected for biopsy when IgA tTG or EMA were positive' **Overall risk of bias**
Number of patients	Total N=268 adults and children
Patient characteristics	Inclusion: adults and children with symptoms suggestive of CD on a gluten containing diet. All patients who received a jejunal biopsy and serology testing were included Exclusion: participants who were IgA deficient
Intervention	All samples analysed by fully automated fluoroenzyme immunoassay tests (Elia Celikey IgA, Elia Gliadin IgA, Elia Gliadin IgG, Elia Gliadin DGP IgA, Elia Gliadin DGP IgG) EMA was analysed by indirect immunofluorescence on monkey oesophagus sections Optimal cut-off values calculated using ROC curves. For all samples, except IgA DGP, best sensitivity and specificity were found to be consisted with manufacturer's recommendations.

Bibliographic reference	Burgin-Wolff (2013): Intestinal biopsy is not always required to diagnose coeliac disease: a retrospective analysis of combined antibody tests
	I. IgA tTG – cut off =7 II. IgA AGA – cut off = 7 III. IgG AGA – cut off = 7 IV. IgA DGP – cut off =7 [a] V. IgG DGP – cut off = 10 VI. IgA EMA – cut off = serum dilution 1:5
Comparison	Jejunal biopsy – no other information. [b]
Length of follow up	
Location	Switzerland
Outcomes measures and effect size	Diagnostic accuracy of serological tests CI: 95%
Source of funding	
Comments	

[a] Authors found a cut off value of 7 instead of 10 increased sensitivity from 71% to 78%, while maintaining specificity
[b] For an undefined period, patients were sometimes selected for biopsy when IgA tTG or EMA were positive

Total N= 149/268 (56%) diagnosed with CD according to Marsh Classification (Marsh 3a, b, or c lesions accepted).

Sensitivity, specificity, PPV, NPV, and efficiency, and positive and negative likelihood ratio data were provided in the paper.

IgA DGP in children and adults
Sensitivity 78% (71 – 85), specificity 97 % (93 – 99), PPV 97% (93 – 99), NPV 78% (71 – 84), efficiency 86%, + LHR = 23. – LHR = 0.23

(a)TP 116	(b)FP 4
(d)FN 33	(c)TN 115

IgG DGP in children and adults
Sensitivity 85% (80 – 90) specificity 92% (86 – 97), PPV 93% (88 – 97), NPV 83% (77 – 90), efficiency 88%, +LHR = 10, -LHR = 0.16

(a)TP 127	(b)FP 10

(d)FN 22	(c)TN 109

IgA tTG in children and adults
Sensitivity 97% (94 – 99) specificity 87% (80 – 92), PPV 90% (85 – 95), NPV 95% (91 – 99), efficiency 92%, +LHR = 7, -LHR = 0.04

(a)TP 144	(b)FP 16
(d)FN 5	(c)TN 103

IgA EMA in children and adults
Sensitivity 98% (96 – 100), specificity 85% (78 – 91), PPV 89% (84 – 94), NPV 97% (94 – 100), efficiency 98%, +LHR = 6, -LHR = 0.02

(a)TP 146	(b)FP 18
(d)FN 3	(c)TN 101

Combinations Tests: 'test combinations containing only IgA antibodies were not considered; they are unsuitable for diagnostic purposes, because of the possibility that some patients may be deficient in IgA'.

**NB: sensitivity and specificity presented here were calculated from raw data values. These differ from the sensitivity and specificity results presented in the paper. The paper presents 'non-classified' data, which relates to the number of patients per test combination that were unable to be classified due to inconsistency between two or more tests (i.e. positive result on one test and negative result in another test(s)). This 'non-classifiable' data was incorporated into the 2x2 tables presented here as false negative data, as it is assumed that the 'non-classified' data was classed as a negative.

Combination of two tests:

IgG DGP + IgA tTG in children and adults
Sensitivity 72% (65 – 80), specificity 96% (92 – 99), PPV 96% (92 – 99), NPV 71% (64 – 79)

(a)TP 108	(b)FP 5
(d)FN 41	(c)TN 114

IgG DGP + EMA in children and adults
Sensitivity 73% (66-80), specificity 95% (91 – 98), PPV 95% (90 – 99), NPV 74% (67-81)

(a)TP 109	(b)FP 6

(d)FN 40	(c)TN 113

IgA DGP + IgG DGP in children and adults
Sensitivity 65% (57 – 72), specificity 99% (98 – 100), PPV 99% (97 – 100), NPV 69% (62 – 76)

(a)TP 97	(b)FP 1
(d)FN 52	(c)TN 118

Combination of three tests:

IgA DGP + IgG DGP + IgA tTG in children and adults
Sensitivity 73% (66-80) specificity 99% (98 – 100), PPV 99% (97 – 100), NPV 75% (68 – 81)

(a)TP 109	(b)FP 1	

(d)FN 40	(c)TN 118	
149	119	

IgA DGP + IgG DGP + EMA in children and adults
Sensitivity 58% (50 – 66), specificity 99 (98 – 100), PPV 99% (96 – 100), NPV 65% (58 – 72)

(a)TP 86	(b)FP 1	
(d)FN 63	(c)TN 118	
149	119	

IgG DGP + EMA + IgA tTG
Sensitivity 70% (62 – 77), specificity 96% (92 – 99), PPV 95% (91 – 99), NPV 73% (66-79)

(a)TP 104	(b)FP 5	

(d)FN 45	(c)TN 114	
149	119	

Combination of 4 tests:
IgG DGP + IgA DGP + EMA + IgA tTG
Sensitivity 56% (48 – 64), specificity 99% (98 – 100), PPV 99% (97 – 100), NPV 64% (58 – 71)

(a)TP 84	(b)FP 1	
(d)FN 65	(c)TN 118	
149	119	

Bibliographic reference	**Clouzeau-Girard (2011): HLA-DQ genotyping combined with serological markers for the diagnosis of celiac disease: Is intestinal biopsy still mandatory?** **Reference ID:**
Study type	Cohort (prospective)
Study quality	1. Could the selection of patients have introduced bias? NO – all patients were consecutively recruited for suspicion of CD

	2. Is there concern that the included patients do not match the review question? No – all patients were suspected of CD 3. Could the conduct or interpretation of the index test have introduced bias? NO – serological testing was carried out according to manufacturer recommendations 4. Is there concern that the index test, its conduct, or interpretation differ from the review question? NO – genotyping and serological testing were as specified in study protocol 5. Could the reference standard, its conduct, or its interpretation have introduced bias? NO – 3 biopsies were taken from the duodenum and classified according to Marsh criteria 6. Is there concern that the target condition as defined by the reference standard does not match the review question? NO – target condition matches that specified in protocol 7. Could the patient flow have introduced bias? NO – all patients consecutively recruited, all received both index and reference tests **Overall risk of bias: Low** – All patients met target population as defined by protocol, had both serology and biopsy, and the index test and reference standard were appropriate to study protocol.
Number of patients	Total N = 170 children
Patient characteristics	Inclusion criteria: 170 patients who underwent serologic testing for coeliac disease and a small bowel biopsy between 2003 and 2006 to investigate chronic symptoms suggestive of CD. Exclusion: children excluded from the study if they had already begun gluten free diet, or if histological examination was inconclusive due to poor orientation of the sample, or if IgA deficiency was found. Patients were classified into two groups:

	Group 1: children with histology suggestive of CD (Marsh-Oberhuber classification 3a-c) Group 2: children with histology not suggestive of CD (histology showing partial villous atrophy but normal intraepithelial lymphocytes)
Intervention	I. HLA DQ2/8 genotyping – PCR performed with specific DQa1 and DQB1 primers. During the study this technique replaced with an allelic typing of the DQB1 gene, and when the susceptibility DQB1 alleles were identified, the DQA1 gene was studied. Both strategies were shown to provide identical results. Results given as positive or negative for the distinct predisposition alleles to determine the presence or absence of these genotypes. II. IgA EMA – determined using indirect immunofluorescence. Considered positive when IgA EMA antibodies were positive according to manufacturer's cut-off values III. IgA TTG – determined using ELISA assay. Considered positive when IgA tTG antibodies were positive according to manufacturer's cut-off values IV. Total IgA – Total IgA was determined in the serum to rule out selective IgA deficiency.
Comparison	Biopsy: two or three biopsies were obtained in the third part of the duodenum during endoscopy.
Length of follow up	
Location	France
Outcomes measures and effect size	Diagnostic accuracy of combined serology and genotyping 95% CI
Source of funding	Not stated
Comments	

82/162 (49%) considered positive for CD according to Marsh-Oberhuber classification (Marsh grade 3a-c).

8/170 excluded (4%): 2 children were already consuming a GFD; 1 child had previously been on a GFD and reintroduced gluten only 8 weeks earlier; 2 children had selective IgA deficiency; 3 children had intestinal biopsies which could not be classified because of bad orientation of the sample.

Of the 82 CD-positive children, 70 carried the DQ2 heterodimer and 6 possessed the DQ8 genotype. 5 patients carried both.

The most common diagnosis of those in the control group included: gastrooesophageal reflux (12.5%); psychological eating disorders (10%); lactose allergy

(2.25%); Iron-deficient anaemia (3.7%); helicobactor pyliori gastritis (10%); inflammatory bowel disease (3.7%); no aetiology (27%)

HLA DQ2/DQ8 genotyping in children

Sensitivity 99% (96 – 100), specificity 69% (59 – 79), PPV 76% (68 – 85), NPV 98% (95 – 100)

(a)TP 81	(b)FP 25
(d)FN 1	(c)TN 55

Combined IgA TTG / IgA EMA and HLA DQ2/DQ8 genotyping in children
Sensitivity 99% (96 – 100), specificity 96% (92 – 100), PPV 96% (92 – 100), NPV 99% (96 -100)

(a)TP 81	(b)FP 3
(d)FN 1	(c)TN 77

Bibliographic reference	**Porcelli (2011): Assessment of combination screening assay for celiac disease**
Study type	Case-control

Study quality	1. Could the selection of patients have introduced bias? Yes – unclear if CD confirmed patients were consecutively recruited. Control population consisted of disease controls with various other conditions and healthy blood donors. It is unclear how control cases were chosen 2. Is there concern that the included patients do not match the review question? No – all patients were confirmed of CD according to histological and serological criteria 3. Could the conduct or interpretation of the index test have introduced bias? No – assays were conducted using cut offs recommendd 4. Is there concern that the index test, its conduct, or interpretation differ from the review question? 5. Could the reference standard, its conduct, or its interpretation have introduced bias? 6. Is there concern that the target condition as defined by the reference standard does not match the review question? 7. Could the patient flow have introduced bias?
Number of patients	Total N = 201 (41 CD patients and 169 control subjects)
Patient characteristics	CD patients: 41 recently diagnosed CD patients (according to histological and serological criteria), mean age 38 years. Controls: N = 169; n=145 'disease controls'; 15 with autoimmune hepatopathies; 12 with cirrhosis; 35 with viral hepatitis; 83 with other gastrointestinal diseases, and n=24 'healthy' blood donors.
Intervention	I. IgA tTG: ELISA (Quanta-Lite human recombinant tTG (h-tTG IgA). manufacturer cut-off. II. IgA Ttg: ELISA (Quanta-Lite human recombinant tTG (h-tTG IgG). manufacturer cut-off. III. IgA DGP: ELISA (Quanta-Lite gliadin IgA II). manufacturer cut-off. IV. IgG DGP: ELISA (Quanta-Lite gliadin IgG II). manufacturer cut-off. V. IgA EMA: Immunoflourescence (Eurospital). manufacturer cut-off.

	VI. IgA and IgG for tTG and DGP in a single assay (QUANTA Lite h-tTG/DGP screen ELISA assay (using purified synthetic DGP's and native human tissue transglutaminase). manufacturer cut-off.
Comparison	Biopsy-confirmed CD, or non-CD
Length of follow up	
Location	Italy
Outcomes measures and effect size	Diagnostic accuracy of serological tests 95% CI
Source of funding	
Comments	

At the time of diagnosis, all CD-confirmed patients had histological signs of Marsh 3a-c.

For the purposes of this analysis, sensitivity, specificity, PPV, and NPV values are derived from comparing CD patients to CD-negative disease controls

IgA + IgG h-tTG/DGP in adults
Sensitivity 100% (100 – 100), specificity 90 (86 – 95), PPV 75% (63 – 86), NPV 100% (100 – 100)

(a)TP	(b)FP
41	14

(d)FN 0	(c)TN 131

Bibliographic reference	**Mubarak (2011): Immunoglobin G antibodies against deamidated gliadin peptides outperform anti endomysium and tissue transglutaminase antibodies in children <2** **Reference ID:**
Study type	Cohort (prospective)
Study quality	1. Could the selection of patients have introduced bias? NO 2. Is there concern that the included patients do not match the review question? NO 3. Could the conduct or interpretation of the index test have introduced bias? NO 4. Is there concern that the index test, its conduct, or interpretation differ from the review question? NO 5. Could the reference standard, its conduct, or its interpretation have introduced bias? NO 6. Is there concern that the target condition as defined by the reference standard does not match the review question? NO 7. Could the patient flow have introduced bias? YES – any subject with abnormal serology was biopsied **Overall risk of bias: LOW**. Study participants met protocol criteria, all underwent index tests and reference standard as stipulated by protocol , reference standard could not have introduced bias, and patient flow was unbiased. However, only participants with abnormal serology was biopsied, which may bias the outcome of establishing accuracy of serology.
Number of patients	N= 212 children suspected of CD ; <2 yrs n=41. Age range 0.6mnts – 17.8 yrs, mean age = 6.3 yrs.
Patient characteristics	Inclusion: Children <2 suspected of having CD in whom both a small intestine biopsy had been done and serological testing (EMA and/or tTGA) in the period 1998-2009. Any patient with abnormal serology, and also patients with negative serology and a high-suspicion of CD were biopsied

	All patients were on GCD, had IgA of at least 0.08 g/L, and did not suffer from giardiasis. Exclusion: none listed
Intervention	IgA DGP and IgG DGP determined using 2 methods: 1) Bindazyme Human anti gliadin EIA Kits IgA and IgG 2) Quanta Lite Gliadin IgA II and IgG II Cut-off ≥ 20 U/mL considered positive Quanta-Lite-kit combined kit used for detection of IgA and IgG-DGP, as well as IgA and IgG tTGA in human serum with a cut-off value of ≥20 U/mL Serum IgA tTGA measured by ELISA using human recombinant tTG. Cut off ≥ 10 U/mL were considered positive
Comparison	Intestinal biopsy [c] Mean of 3.2 biopsies per patient taken from distal duodenum by upper endoscopy. All biopsies revised by single pathologist . Histological diagnosis of CD made using Marsh modified classification. Marsh I and Marsh II were regarded as not conclusive for CD. Marsh III villous atrophy considered diagnostic for CD
Length of follow up	
Location	
Outcomes measures and effect size	diagnostic accuracy of serological tests 95% CI

[c] The pathologist was blinded to clinical presentation and serological results.

Source of funding	Not stated
Comments	

Total N =109/ 212 (total - 51.4%, children >2 = 83/171 Children <2 – 26/41 = 46.4%) diagnosed with CD using Marsh criteria – Marsh III lesion considered CD positive.

2 of remaining CD negative participants had a Marsh I lesion
1 of remaining CD negative participants had Marsh II lesion

PPV, NPV and all confidence intervals presented in the paper in a 2 x 2

a-DGP/tTGA children ≥2 yrs
sensitivity 98% (91-100) specificity 56% (45-66) PPV 68% (58-76) NPV 96% (85-99)

(a)TP 81	(b)FP 39
(d)FN 2	(c)TN 49

a-DGP/tTGA children < 2 yrs
Sensitivity 100% (84-100) specificity 93% (66-100) PPV 96% (79-100) NPV 100% (73-100)

(a)TP 81	(b)FP 39
(d)FN 2	(c)TN 49

Bibliographic reference	**Swallow (2013): Quality not quantity for transglutaminase antibody 2: the performance of an endomysial and tissue transglutaminase test in screening coeliac disease remains stable over time**
Study type	Cohort (retrospective)
Study quality	1. Could the selection of patients have introduced bias? NO – all patients were recruited for suspicion of CD 2. Is there concern that the included patients do not match the review question? No – all patients were suspected of CD 3. Could the conduct or interpretation of the index test have introduced bias? NO – serological testing was carried out according to manufacturer recommendations 4. Is there concern that the index test, its conduct, or interpretation differ from the review question? NO – genotyping and serological testing were as specified in study protocol 5. Could the reference standard, its conduct, or its interpretation have introduced bias? NO – 3 biopsies were taken from the duodenum and classified according to Marsh criteria 6. Is there concern that the target condition as defined by the reference standard does not match the

	review question? NO – target condition matches that specified in protocol 7. Could the patient flow have introduced bias? NO – authors have confirmed that all 756 participants received biopsy, IgA tTG, and IgA EMA. **Overall risk of bias: Low:** Patient selection, population, index test, comparator, and target condition all match protocol outline. It is unclear, however, if the decision to biopsy was driven by serological results, and therefore, if all patients underwent serological testing and biopsy.
Number of patients	Total N = 756
Patient characteristics	Inclusion criteria: all new patients seen between 2008 – 2009 who had been tested for tTG and EMA and had a duodenal biopsy performed. All patients were on a gluten containing diet at the time of biopsy. Exclusion criteria: Patients were excluded if only one serological test was done, or if serological testing was not carried out within 12 weeks of biopsy. Patients being monitored for pre-existing coeliac disease were also excluded from the audit.
Intervention	I. IgA tTG – ELISA test (AUESKULISA). Results interpreted as negative if < 15 U/ml, equivocal 15-5- U/ml, or positive <50 U/ml. All units are arbitrary and ssay-specific; there is no international standard to ensure comparability between assays . 2 levels of internal quality control material (IQC) with equivocal and positive results and kkit controls are assayed on each run to assess the validity of the results. II. IgA EMA – assessed by indirect immunofluorescence on monkey oesophagus tissue. Interpreted as negative, weal positive, positive, or strong positive. Weak positive and negative EMA internal quality control materials and regular review of consistency of reading thresholds are used to maintain stable reporting practice and assay sensitivity over time.
Comparison	Duodenal biopsy (Marsh grade 3 taken as CD)
Length of follow up	
Location	UK
Outcomes measures and effect size	Diagnostic accuracy of serological test 95% CI
Source of funding	None declared
Comments	

23/756 (3.04%) patients positive for CD according to Marsh grade 3. (730 controls)

Marsh grades 1 -3 lesions were found in 30 patients. Results presented here are based only on Marsh 3 lesions.

Data compared for 04 – 06 data (Hopper 08 paper) and 08 – 09 data in order to examine whether data reproducible. 08 – 09 presented here, 04 – 06 data presented in Hopper paper.

2 step strategy: TtG positive OR equivocal , then EMA positive
Sensitivity 87% (65 – 97), specificity 97% (95 – 98),

(a)TP 20	(b)FP 23
(d)FN 710	(c)TN 3

2 step strategy: TtG and EMA positive
Sensitivity 83% (60 – 94), specificity 99% (98 – 99.6),

(a)TP 19	(b)FP 7

(d)FN 4	(c)TN 726

Reference	Study type/ Evidence level	Number of patients	Patient characteristics	Intervention	Comparison	Length of follow-up	Outcome measures	Source of funding
Hopper AD, Hadjivassiliou M, Hurlstone DP, Lobo AJ, McAlindon ME, Egner W *et al*. What is the role of serologic testing in celiac disease? A prospective, biopsy-confirmed study with economic	Cohort/case control	N=2000 Adults (≥16yrs) UK	Inclusion: consecutive adults referred for gastroscopy without a previous diagnosis of celiac disease at a single endoscopist department from January 2004 to April 2006, N=1167 (58.3%) female, mean age 55.8yrs (range 16 to 94yrs) Exclusion: known diagnosis of coeliac disease, a coagulopathy (international normalised ratio > 1.3 or platelet count of < 80), active GI bleed or a suspected carcinoma observed during the examination (group 2: patients with a known diagnosis of celiac disease on a GFD for >1yr undergoing repeat duodenal biopsies and	IgA/IgG AGA (ELISA, AESKU Diagnostics)(cut-off > 15 U/mL) IgA tTG (ELISA, AESKU Diagnostics)(cut-off > 15 U/mL) IgA EMA (immunofluorescence, primate oesophagus)	Policy of 4 duodenal biopsy specimens from the second part of the duodenum		Marsh criteria Those with villous atrophy with supporting signs and symptoms were considered to have coeliac disease Those with villous atrophy (confirmed on a second review of the sample to ensure a well-oriented sample) and a antibody –	Not stated

analysis. *Clinical Gastroenterology & Hepatology* 2008;**6**:314-20			serologic analysis – results not included in this table)	Total IgA (Behring BN2 nephelometer, Siemens) Blood for serological tests taken at the same time as the biopsy			ve profile were classed as seronegative coeliac disease, to confirm this they were required to have DQ2 or DQ* pattern consistent with CD and a clinical and histological response to a GFD	
Effect size: (CI 95%) N=77/1000 diagnosed with coeliac disease (prevalence, all patients attending for gastroscopy of 3.9%); N=29 Marsh 3a, N=30 Marsh 3b, 18 Marsh 3c lesions IgA deficiency 0.7% (N=14/2000) Symptoms (coeliac disease vs. non coeliac disease): Weight loss (15.6% vs. 5.3%), p<0.05 Diarrhoea (42.9% vs. 5.2%), p<0.05 Dyspepsia (17.3% vs. 1%), p<0.05								

Reflux (13.8% vs 1%), p<0.05

Dysphagia (7.2% vs. 0%), p<0.05

Those with coeliac disease were significantly younger (mean age 48.0 vs. 56.1 yrs), p<0.05, there were significantly more females (70.1% vs. 57.9%), p<0.05, than those without coeliac disease

IgA tTG

sensitivity 90.9% (82.4 to 94.5), specificity 90.9% (89.5 to 92.1), PPV 28.6% (23.3 to 34.5), NPV 99.6% (99.2 to 99.8)

2X2:

(a) TP 70	(b) FP 175
(c) FN 7	(d) TN 1748

IgA EMA

sensitivity 87.0% (77.7 to 92.8), specificity 98.0% (97.4 to 98.6), PPV 64.4% (54.9 to 73.0), NPV 99.4% (99.0 to 99.7)

2X2:

(a) TP 67	(b) FP 37
(c) FN 10	(d) TN 1886

If tTG +ve and then EMA +ve (2-step)

sensitivity 85.7% (76.2 to 91.8), specificity 98.6% (98.0 to 99.0), PPV 71.7% (61.8 to 79.9), NPV 99.4% (99.4 to 99.0)

2X2:

(a) TP 66	(b) FP 26
(c) FN 11	(d) TN 1897

Both tTG +ve and EMA +ve

sensitivity 85.7% (76.2 to 91.8), specificity 98.6% (98.0 to 99.0), PPV 71.7% (61.8 to 79.9), NPV 99.4% (99.4 to 99.0)

2X2:

(a) TP 66	(b) FP 26
(c) FN 11	(d) TN 1897

Either tTG +ve or EMA +ve

sensitivity 92.2% (84.0 to 96.4), specificity 90.3% (88.9 to 91.6), PPV 27.6% (22.5 to 33.4), NPV 99.7% (99.3 to 99.8)

2X2:

(a) TP 71	(b) FP 186
(c) FN 6	(d) TN 1737

IgG AGA

sensitivity 48.1% (37.3 to 59.0), specificity 95.8% (94.9 to 99.6), PPV 31.6% (23.9 to 40.5), NPV 97.9% (97.1 to 98.4)

2X2:

(a) TP 37	(b) FP 77
(c) FN 40	(d) TN 1849

IgA AGA

sensitivity 49.4% (38.5 to 60.2), specificity 89.6% (88.2 to 90.1), PPV 16.0% (11.9 to 21.2), NPV 97.8% (97.0 to 98.4)

2X2:

(a) TP 38	(b) FP 200
(c) FN 39	(d) TN 1723

Both IgA and IgG AGA

sensitivity 36.4% (26.5 to 47.5), specificity 98.8% (98.2 to 99.2), PPV 54.9% (41.4 to 67.7), NPV 97.4% (96.7 to 98.1)

2X2:

(a) TP 28	(b) FP 23
(c) FN 49	(d) TN 1900

Using only IgA tTG +ve 245 would have undergone biopsy and 1 in 11 cases of coeliac disease would have been missed

Using only IgA EMA +ve 104 would have undergone biopsy and 1 in 8 cases of coeliac disease would have been missed

Using IgA tTG +ve and then IgA EMA +ve 92 would have undergone biopsy and 1 in 7 cases of coeliac disease would have been missed

Using either IgA tTG +ve or IgA EMA +ve 257 would have undergone biopsy and 1 in 13 cases of coeliac disease would have been missed

Those with partial villous atrophy (Marsh 3a or 3b) had significantly lower mean tTG titre (168.1 U/mL and 165.0 U/mL) than those with total villous atrophy (255 U/mL), $p<0.05$

Those with Marsh 1 or 2 had significantly lower mean tTG titre (27.7 U/mL and 23.0 U/mL) than those with villous atrophy, $p<0.05$

EMA sensitivity 79% in partial atrophy, 100% in total atrophy, p<0.01

tTG sensitivity 86.0% (Marsh 3a), 100% (Marsh 3c), p<0.05

QUADAS:

1. Could the selection of patients have introduced bias? NO - All patients were consecutively recruited
2. Is there concern that the included patients do not match the review question? NO – patients matched review protocol
3. Could the conduct or interpretation of the index test have introduced bias? NO – manufacturer test cut-off used.
4. Is there concern that the index test, its conduct, or interpretation differ from the review question? NO
5. Could the reference standard, its conduct, or its interpretation have introduced bias? NO – reference standard matched review protocol
6. Is there concern that the target condition as defined by the reference standard does not match the review question? NO – target condition matched review protocol
7. Could the patient flow have introduced bias? NO – all patients received same reference standard and were included in analyses in accordance with review protocol

Overall risk of bias: LOW – Patients Were consecutively recruited. Index and reference tests, and target condition matched review protocol. All participants received the same reference standard and index tests.