

High-titre circulating tissue transglutaminase-2 antibodies predict small bowel villous atrophy, but decision cut-off limits must be locally validated

L. Beltran,^{*} M. Koenig,[†] W. Egner,[‡]
M. Howard,[†] A. Butt,[§] M. R. Austin,[§]
D. Patel,[‡] R. R. Sanderson,[‡]
S. Goubet,^{**} F. Saleh,^{*} J. Lavender,^{*}
E. Stainer^{*} and M. D. Tarzi^{*††}

Departments of ^{*}Biochemistry and Immunology,
[†]Histopathology and [§]Gastroenterology, ^{**}Clinical
Investigation and Research Unit, Royal Sussex
County Hospital, [§]Department of Paediatric
Gastroenterology and Nutrition, Royal Alexandra
Children's Hospital, Brighton and Sussex
University Hospitals NHS Trust, ^{††}Brighton and
Sussex Medical School, University of Sussex,
Brighton, and [‡]Department of Immunology
Protein Reference Unit and UK NEQAS for
Immunology, Immunochemistry & Allergy,
Northern General Hospital, Sheffield, UK

Accepted for publication 5 December 2013
Correspondence: L. Beltran, Department of
Biochemistry and Immunology, Royal Sussex
County Hospital, Brighton and Sussex
University Hospitals NHS Trust, Eastern Road,
Brighton BN2 5BE, UK.
E-mail: luisa.beltran@nhs.net

Introduction

The detection of circulating immunoglobulin (Ig)A tissue transglutaminase-2 (TTG2) antibodies is a standard first-line investigation for coeliac disease (CD) [1]. Current UK guidelines [2] recommend the investigation of seropositive individuals by histopathological examination of multiple duodenal biopsy samples, scored against the histological criteria proposed by Marsh [3,4].

Several studies report excellent predictive value for CD or high-risk Marsh 3 histology at high TTG2 antibody titres [5–11], questioning the requirement for routine duodenal biopsy in this setting. However, allocating a TTG2 antibody

Summary

Numerous studies suggest that high levels of circulating immunoglobulin (Ig)A tissue transglutaminase (TTG2) antibodies predict coeliac disease with high specificity. Accordingly, it has been suggested that duodenal biopsy may not be required routinely for diagnostic confirmation where quantitative serology identifies the presence of high antibody titres. However, defining a cut-off TTG2 threshold is problematic, as the multiple available assay methods are not harmonized and most studies have been focused on the paediatric population. Recent paediatric guidelines proposed a TTG2 antibody diagnostic cut-off at 10 × the upper limit of normal (ULN) for the method; however, concerns remain about errors of generalization, between both methods and laboratories. In this study, we used retrospective laboratory data to investigate the relationship between TTG2 antibody levels and Marsh 3 histology in the seropositive population of adults and children at a single centre. Among 202 seropositive patients with corresponding biopsies, it was possible to define a TTG2 antibody cut-off with 100% specificity for Marsh 3 histology, at just over 10 × ULN for the method. However, UK National External Quality Assurance Scheme returns during the study period showed a wide dispersion of results and poor consensus, both between methods and between laboratories using the same method. Our results support the view that high-titre TTG2 antibody levels have strong predictive value for villous atrophy in adults and children, but suggest that decision cut-offs to guide biopsy requirement will require local validation. TTG2 antibody assay harmonization is a priority, in order to meet the evolving requirements of laboratory users in this field.

Keywords: coeliac disease, endomysial antibody, TTG2 antibody

decision cut-off for this purpose is problematic: in the absence of an international reference preparation, TTG2 antibody units and reference ranges are arbitrary and method-specific. Recent European Society of Paediatric Gastroenterology, Hepatology and Nutrition (ESPGHAN) consensus guidelines propose normalizing TTG2 antibody results to multiples of the upper limit of normal (ULN) [12]. According to the guidelines, biopsy is potentially avoidable in symptomatic children with TTG2 antibody levels in excess of 10 × ULN, provided that other criteria are also fulfilled. The option of this approach has also been adopted in recent UK guidelines developed by the Coeliac Disease Working Group of The British Society of Paediatric

Gastroenterology, Hepatology and Nutrition (BSPGHAN), in collaboration with Coeliac UK [13].

From a laboratory perspective, such an approach raises important questions of quality and achievability. First, it is not clear that normalization to ULN truly harmonizes results between the myriad commercially available detection systems [14,15], all of which report arbitrary units with method-specific reference ranges. This is a particularly important issue, as much of the published data relate to a single manufacturer's methods. Secondly, it is not clear that the reproducibility of results between centres is sufficiently robust to support such guidance, even when the centres are using identical methodology [14,15].

The objectives of this study were to: (i) explore the performance characteristics of a popular but less intensively studied enzyme-linked immunosorbent assay (ELISA) method for TTG2 antibody detection, using retrospective laboratory data to relate TTG2 antibody level to Marsh 3 histology; (ii) define a cut-off TTG2 antibody level with high specificity for Marsh 3 histology; and (iii) explore the implications of applying such cut-offs between different centres.

Materials and methods

Study group

The laboratory information management system was interrogated for all positive TTG2 antibody requests received between August 2010 and January 2013 that had corresponding duodenal biopsy reports. In this way, an anonymized database was populated with TTG2 antibody values, Marsh histology scores and associated demographic/clinical information. Selected serology–biopsy pairs were then excluded on the following grounds: outdated (more than 120 days between serology and biopsy, or more than 30 days between biopsy and serology where biopsy was performed first); biopsy indication to monitor dietary compliance in a known CD patient; technically inadequate biopsy according to the reporting histopathologist; or the primary presentation was dermatitis herpetiformis. The Research and Development Department at Brighton and Sussex University Hospitals NHS Trust reviewed the protocol and confirmed that it fulfilled the criteria for service development, as we aimed to evaluate the implementation of published guidelines using existing anonymized laboratory data; accordingly, ethical approval was not required.

Serological and histological analysis

TTG2 antibody levels were measured by immunoassay (QuantaLite R h-tTG IgA ELISA; Inova Diagnostics Inc., San Diego, CA, USA) using a DS2 automated ELISA workstation. Results were reported in U/ml using the

manufacturer's supplied reference ranges: 0–4 negative; 4–10 weak positive; ≥ 10 positive. The highest standard in the calibration curve for this assay is 100 U/ml. Sera with optical density values exceeding that of this top standard were generally reported as >100 U/ml, but occasionally a quantitative result of more than 100 was reported when the optical density of the sample lay between that of the top standard and the upper optical density limit of the plate reader. A kit-independent internal quality control sample was analysed with each assay run and out-of-control runs [as judged by the presence of a quality control value outside the previously determined mean $\pm 2 \times$ standard deviation (s.d.) range] were rejected.

Duodenal biopsy histology was reported in the context of a routine clinical service, using routine specimen processing and haematoxylin and eosin staining. Accordingly, the histopathologists were not blinded to the serological status and clinical details of the patients.

UK NEQAS data

Participation in external quality assurance schemes is a requirement for all accredited diagnostic laboratories. The UK National External Quality Assurance Scheme (UK NEQAS) regularly distributes identical samples to participants in the CD serology scheme and collates the returns, providing a rich data source to investigate assay reproducibility. Anonymized UK NEQAS CD serology scheme returns from three different distributions were populated in a database. We interpreted the returned quantitative results according to the manufacturer's supplied reference ranges. The qualitative interpretation provided by the participating laboratory was disregarded. To enable the comparison of quantitative data between methods, TTG2 antibody levels were normalized by expression as a multiple of the ULN for the method, defined as the lowest value supplied in the assay datasheet that could be considered abnormal. The methods included in this study with associated manufacturer reference ranges and limits applied are detailed in Table 1.

Statistical analysis

Primary data were populated in Microsoft Excel. Results were exported to Graphpad Prism for statistical analysis. A precision-based method was utilized to explore the predictive value of TTG2 quantification for Marsh 3 histology. A construct set was generated by randomly selecting half the data set, using a random number generator (<http://www.random.org>), to generate a receiver operating characteristics (ROC) curve and to define a cut-off TTG2 antibody value with high specificity for Marsh 3 histology. The remaining data set formed the validation set and was used to validate the previously defined cut-off.

Table 1. Tissue transglutaminase (TTG)2 detection systems studied in this project, with manufacturer-supplied reference ranges and limits applied in this study.

Method	Supplied reference ranges (U/ml unless stated otherwise)	Upper limit of normal applied in this study (U/ml unless stated otherwise)	
Inova	0–4	Negative	4
Quantalite	4–10	Weak positive	
	>10	Positive	
Phadia 250	<7	Negative	7
ELIA	7–10	Equivocal	
	>10	Positive	
Orgentec	<10	Negative	10
Euroimmun	<20 RU/ml	Negative	20 RU/ml
Phadia	<5	Negative	5
Varelista	5–8	Equivocal	
	>8	Positive	
Aesku	<15	Negative	15

Results

TTG2 antibody assay characteristics

The mean coefficient of variation for the assay during the study period was 10.3%, based on serial analysis of kit-independent quality control material with every run. The laboratory maintained good performance throughout on the UK NEQAS CD serology scheme.

All samples analysed using the TTG2 antibody assay were screened for IgA deficiency. Total IgA quantification was performed in parallel with all CD serology requests received between August 2010 and September 2012, identifying 64 individuals with specific IgA deficiency (defined as IgA below the assay detection limit of 0.1 g/l) and 48 patients with more severe degrees of partial IgA deficiency (defined as $\text{IgA} \leq 0.2 \text{ g/l}$). To improve laboratory efficiency, we explored the utility of screening by optical density (OD), whereby low OD results on TTG2 antibody assay may indicate IgA-deficient sera. Over short periods of time, it was possible to define an OD threshold with excellent diagnostic accuracy for IgA deficiency; however, the OD cut-off value did not remain stable, being sensitive to instrument servic-

ing and changes in kit lot number. An alternative approach of selecting the five sera with lowest OD values from each assay run proved to be more robust: during a 3-month evaluation period, during which total IgA was quantified for all samples ($n = 2290$), this screening approach successfully identified all six patients with complete IgA deficiency and all five samples with IgA of 0.2 g/l or lower. In the 4 months following implementation of this screening approach (3334 requests processed), the detection rate for IgA-deficient sera has remained fairly constant (0.36% before and 0.30% after). No patients with IgA deficiency were included in this study.

Study group characteristics

A total of 245 patients with positive serology and corresponding biopsies were identified between August 2010 and January 2013. Forty-three patients were excluded for the following reasons: technically inadequate biopsy ($n = 4$); outdated ($n = 29$); indication for biopsy was DH ($n = 5$); and indication for biopsy was CD monitoring on gluten-free diet ($n = 5$). The remaining 202 patients comprised the study group, whose characteristics are displayed in Table 2.

High-titre TTG2 levels are highly specific for mucosal villous atrophy

Median TTG2 antibody concentrations were observed to generally increase with increasing Marsh histology stage (Fig. 1). High-titre antibody was a striking feature of the Marsh 3 population, but lower levels of TTG2 antibody were also clearly compatible with villous atrophy.

To determine a cut-off TTG2 antibody level with high specificity for villous atrophy (Marsh histology stage 3), construct and validation sets were generated by randomly selecting half the study group, as detailed previously. The construct and validation sets were well-matched for Marsh histology score, age, gender and TTG2 antibody level (Table 3). A receiver operating characteristic (ROC) curve was generated using the construct set data with area under the curve of 0.81 [95% confidence interval = 0.79–0.90] (Fig. 2). A TTG2 antibody threshold of $\geq 45 \text{ U/ml}$ ($11.25 \times \text{ULN}$ for the assay) was chosen as a cut-off with

Table 2. Characteristics of the study group.

	Entire study group	Adults only (>18 years)	Paediatric only (<18 years)
<i>n</i>	202	166 (82.2%)	36 (17.8%)
Mean age (years)	40	47	8
Age range (years)	1 to 86	18 to 86	1–17
Female: male	129:73 (63.9% female)	100 : 66 (60.2% female)	29 : 7 (80.6% female)
Marsh 0	31	28	3
Marsh 1	7	7	0
Marsh 2	3	2	1
Marsh 3	161	129	32

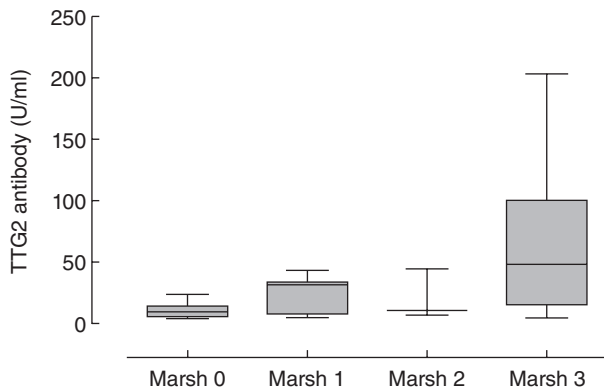


Fig. 1. Box-and-whisker plots of tissue transglutaminase (TTG)2 concentrations related to Marsh score in 202 seropositive individuals with corresponding biopsies received between August 2010 and January 2013 at Brighton and Sussex University Hospitals NHS Trust.

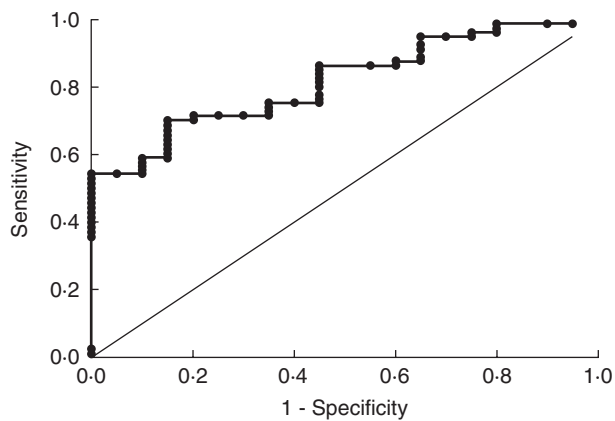


Fig. 2. Receiver operating characteristics (ROC) curve depicting the sensitivity and specificity of tissue transglutaminase (TTG)2 antibody concentration in relation to Marsh 3 histology in the construct set.

100% specificity for Marsh 3 histology (Table 4a), with all study subjects above this threshold reported as Marsh 3. Application of this cut-off to the validation set data also achieved specificity of 100%, thereby validating the cut-off value (Table 4b).

A ROC curve constructed around the entire data set is displayed in Fig. 3. The sensitivity for Marsh 3 histology at a cut-off of 45 U/ml was 51.6% in the entire study group, or

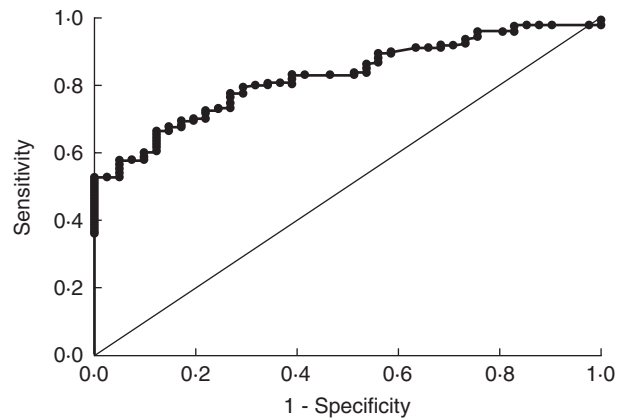


Fig. 3. Receiver operating characteristics (ROC) curve depicting the sensitivity and specificity of tissue transglutaminase (TTG)2 antibody concentration in relation to Marsh 3 histology in the entire study group.

69.4% in the paediatric subpopulation. Applying a more stringent cut-off at 100 U/ml (the top standard for the assay) achieves sensitivity of 35% overall and 65% in the paediatric subgroup.

The cut-off limit cannot be generalized to other centres using the same ELISA system

In order to investigate the suitability of our chosen cut-off for general use among users of the same method, we analysed UK NEQAS CD serology scheme returns from approximately 30 laboratories using the Inova Quantalite ELISA system. Distributions 115 (September 2011), 121 (February 2012) and 122 (April 2012) were selected for analysis, representing increasingly positive samples according to the method-specific, all-laboratory mean. The designated target response in all three distributions was positive.

All users of the Inova methods returned results in the positive range (> 4 U/ml) in all three distributions. However, quantitative results were dispersed very widely within the positive range (Fig. 4). In distributions 115 and 121, there would have been poor agreement between centres as to whether the sample lay above or below our cut-off of 45 U/ml, with approximately half the laboratories on either side of this threshold. Values relating to the very strong

Table 3. Construct set *versus* validation set. Half the study group ($n = 101$) were selected randomly to form the construct set and used to determine a cut-off with high specificity for Marsh 3 histology. The remaining 101 subjects formed the validation set.

	Marsh 3	Marsh 0, 1 or 2	Female (%)	Mean age (years)	Mean TTG2 (U/ml)
Construct set	$n = 81$	$n = 20$	61.3	42	51.6
Validation set	$n = 80$	$n = 21$	65.7	38	45.6

TTG2 = tissue transglutaminase.

Table 4. (a) Results of the application of a tissue transglutaminase (TTG)2 cut-off of 45 U/ml to the construct group

	Marsh 3	Marsh 0, 1 or 2	Total
TTG2 ≥ 45 U/ml	<i>n</i> = 44	<i>n</i> = 0	44
TTG2 < 45 U/ml	<i>n</i> = 37	<i>n</i> = 20	57
Total	81	20	101

(b) Results of the application of a tissue transglutaminase (TTG)2 cut-off of 45 U/ml to the validation group

	Marsh 3	Marsh 0, 1 or 2	Total
TTG2 ≥ 45 U/ml	<i>n</i> = 40	<i>n</i> = 0	40
TTG2 < 45 U/ml	<i>n</i> = 40	<i>n</i> = 21	61
Total	80	21	101

positive sample of distribution 122 were also dispersed widely, but with much better consensus regarding the cut-off at 45 U/ml. Nevertheless, considerable overlap is observed between distributions 115 and 121.

Normalization to ULN does not harmonize TTG2 antibody assays

UK NEQAS results from distributions 115, 121 and 122 for several popular methods were analysed to explore the possibility of generalizing normalized cut-offs, both within and between methods. TTG2 antibody levels were normalized to ULN, as described in the Methods section. A cut-off threshold of 10 × ULN was applied, in accordance with ESPGHAN guidelines.

Combined results from all methods are displayed as a frequency histogram in Fig. 5. According to consensus by all-method, all-laboratory mean, distributions 115, 121 and 122 represented weak positive, positive and strong positive sera at 3, 8.4 and 12.4 × ULN, respectively. A wide dispersion of results is evident across all methods and laboratories, with results from a single distribution ranging from 1.4–45.3 × ULN. Consensus regarding a cut-off at

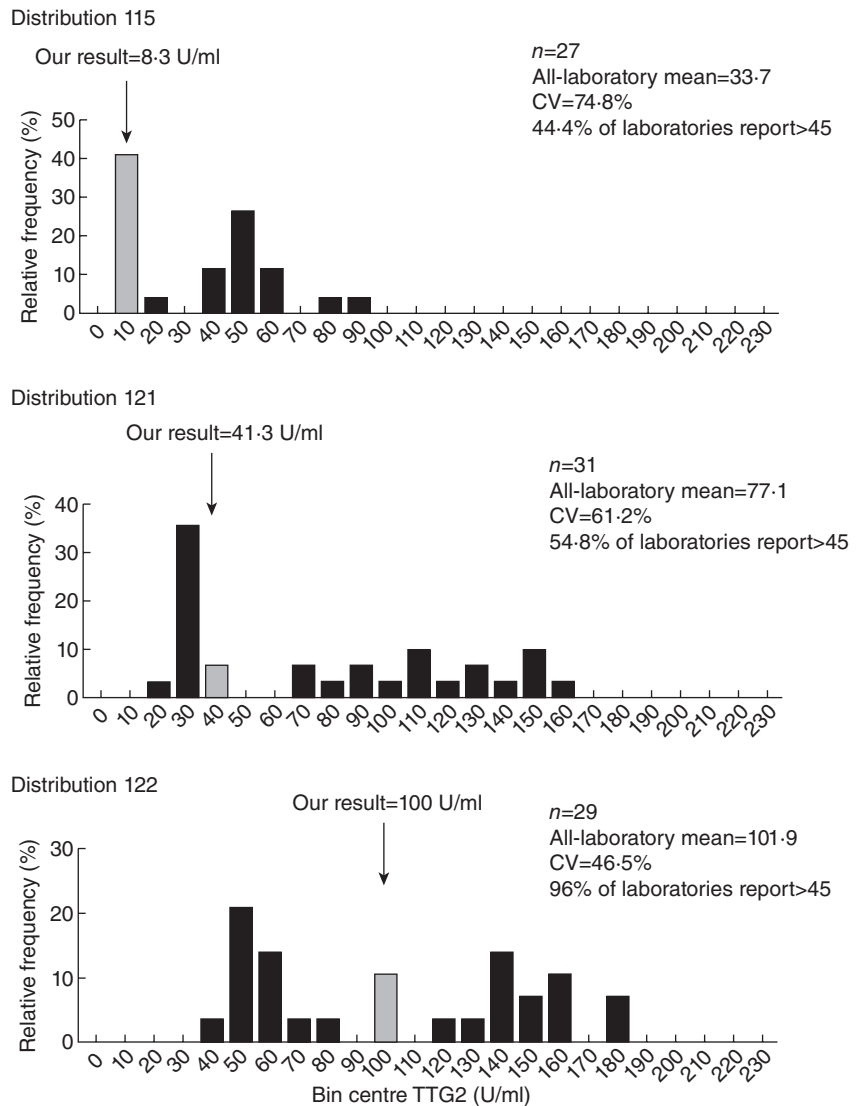


Fig. 4. Frequency histogram plot illustrating the distribution of binned tissue transglutaminase (TTG)2 values returned by laboratories using the Inova method participating in the UK National External Quality Assurance Scheme (UK NEQAS) coeliac disease (CD) serology scheme. Grey columns indicate the bin centre corresponding to the value returned by the Immunology Laboratory at Brighton and Sussex University Hospitals NHS Trust.

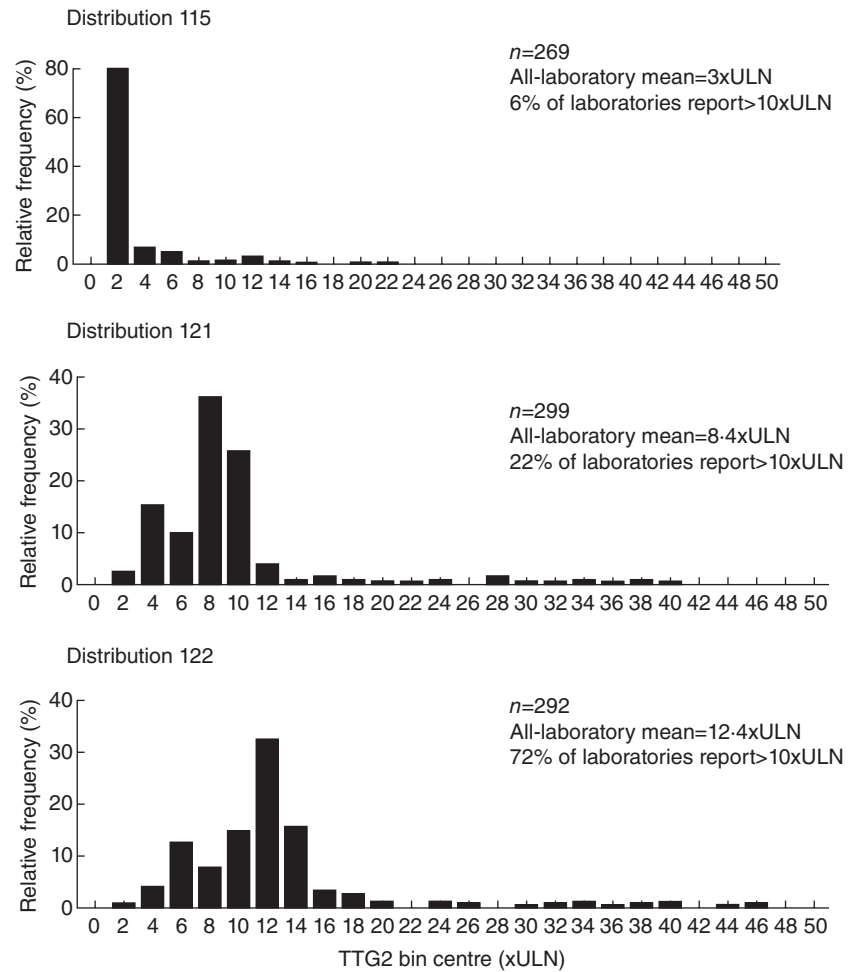


Fig. 5. Frequency histogram plot illustrating the distribution of binned tissue transglutaminase (TTG)2 values returned by participating laboratories in the UK NEQAS CD serology scheme across all methods. TTG2 values have been normalized to multiples of upper limit of normal (ULN).

10 × ULN would have been reasonable in distribution 115, but poor in 121 and 122.

The more detailed method-specific data in Table 5 shows that the poor consensus relates partly to a wide dispersion of results within method groups, but also to marked divergence between the methods when results are expressed as multiples of ULN. Regarding the weak positive sample in distribution 115, Inova users were split fairly evenly above and below 10 × ULN, whereas all other methods returned results below 10 × ULN. The divergence was more marked in distribution 121: Inova users remained fairly evenly split, whereas few users of the Varelisa method and no Orgentec users reported results over 10 × ULN; there was some disagreement between Phadia 250 users, and marked disagreement between Aesku users. In distribution 122, the vast majority of results generated using the Inova, Phadia 250 and Euroimmun methods were in excess of 10 × ULN, in addition to 69.2% of Aesku users; however, no Orgentec users and only 29.2% of Varelisa users reported results above this cut-off.

Discussion

The diagnosis and monitoring of CD have been revolutionized over the last two decades by advances in autoimmune serology [16]. Numerous publications support the view that high levels of circulating TTG2 antibody are specific for CD [5–11], raising the attractive possibility of using cut-off limits to guide the requirement for biopsy [12,13,17]. The first objective of this study was to validate this approach locally in adults and children, using a popular TTG2 antibody detection system that has been studied less intensively to date; the second was to define a local cut-off TTG2 level with high specificity for Marsh score 3 histology; the third was to explore the suitability of the application of such cut-offs between centres.

To determine a cut-off, we utilized retrospective laboratory data to construct a ROC curve relating TTG2 concentration to Marsh 3 histology. There are limitations that constrain the utility of this methodology: Marsh 3 histology has been used as a surrogate marker for CD – a clinical

Table 5. UK National External Quality Assurance Scheme (UK NEQAS) return data from the coeliac disease (CD) serology scheme distributions 115, 121 and 122 across all methods. Tissue transglutaminase (TTG)2 results have been normalized to multiples upper limit of normal (ULN) for each method.

	Method	Valid returns	All-laboratory mean (0.21×ULN)	Lowest (×ULN)	Highest (×ULN)	Proportion > 10 × ULN	Proportion < 10 × ULN
Distribution 115	All methods	269	3.0	1.5	21.8	6%	94%
	Inova	27	8.4	1.5	21.8	55.6%	44.4%
	Phadia 250	154	2.3	1.6	3.4	0.0%	100.0%
	Orgentec	41	1.5	1.1	1.8	0.0%	100.0%
	Euroimmun	17	5.1	3.7	7.1	0.0%	100.0%
	Phadia Varelisa	18	1.5	1.0	2.8	0.0%	100.0%
	Aesku	12	3.9	1.5	7.5	0.0%	100.0%
Distribution 121	All methods	299	8.4	1.6	39.7	22%	78%
	Inova	31	19.3	6.1	39.7	58.1%	41.9%
	Phadia 250	168	8.7	2.8	12.1	14.3%	85.7%
	Orgentec	43	4.7	3.0	7.4	0.0%	100.0%
	Euroimmun	20	10.0	7.3	28.0	95.0%	5.0%
	Phadia Varelisa	23	4.9	2.6	11.4	4.3%	95.7%
	Aesku	14	8.1	1.6	18.7	28.6%	71.4%
Distribution 122	All methods	292	12.4	1.4	45.3	72%	28%
	Inova	29	25.5	8.8	45.3	96.6%	3.4%
	Phadia 250	165	12.1	1.4	17.4	89.7%	10.3%
	Orgentec	42	6.0	3.3	9.4	0.0%	100.0%
	Euroimmun	19	14.0	10.0	44.5	100.0%	0.0%
	Phadia Varelisa	24	8.7	3.4	18.2	29.2%	70.8%
	Aesku	13	13.5	1.9	24.4	69.2%	30.8%

diagnosis would have been preferred and histology itself may produce variable results [18]; validation of the derived cut-off at our centre was performed retrospectively, therefore further prospective evaluation would be required prior to implementation; only the seropositive population has been considered, therefore the sensitivity of TTG2 concentration is assumed to be 100%; the reporting histopathologists were not blinded to the serological and clinical status of the patients; we cannot be completely sure that all patients were taking regular dietary gluten; and the study group had a very high prevalence of Marsh 3 histology and does not represent all populations under investigation for possible CD. We would therefore stress that the first part of this study only really considers whether duodenal biopsy adds extra diagnostic information in the seropositive population referred for biopsy at our centre. It certainly does not provide an estimate of the true diagnostic performance of serology in the general population, which has been recently reviewed elsewhere [16].

With these limitations in mind, we were readily able to define a TTG2 antibody concentration value with high specificity for Marsh 3 histology. Using the Inova Quantalite ELISA method, TTG2 antibody values in excess of 45 U/ml were 100% specific for Marsh 3 histology, with an associated 95% CI between 91–100%. The sensitivity at this cut-off was 51.6%, implying a considerable number of potentially avoidable biopsies. In the paediatric subpopulation the sensitivity was 69.4% – this is lower than

reported in previous studies of children, but the small sample size limits the interpretation of this observation. We also observed that our study population had a higher prevalence of Marsh 3 histology than reported by previous studies [5–8]. A possible explanation is that our study focused on adults, whose characteristics may differ from children; another possible factor is that unblinded histopathology reporting may have affected the distribution of Marsh scores [18], particularly at the Marsh 1/2 borderline. Our derived cut-off for this method was equivalent to 11.25 × ULN and therefore apparently supportive of recent guidelines [12]. However, implementation would depend critically on results being comparable between different laboratories using the same method.

To explore this issue of reproducibility, we analysed data from three different UK NEQAS returns, corresponding (by consensus between all users and all methods) to weak positive, positive and strong positive sera. Users of the Inova method were in full agreement with the positive target response across all three distributions, indicating that the kit is entirely suitable for its intended purpose. However, there was very wide dispersion of quantitative TTG2 results between different laboratories. The application of our cut-off of 45 U/ml would have resulted in poor consensus between centres for the samples with moderately raised TTG2 antibody (distributions 115 and 121), although almost all laboratories would have been in agreement with regard to the sample with very high TTG2 levels (distribu-

tion 122). It is perhaps tempting to propose a cut-off of 100 U/ml (upper limit of the assay equivalent to $25 \times \text{ULN}$) for the method to reduce the margin for error, at the expense of reduced sensitivity to 35% overall and 65% in the paediatric subgroup. However, even at this threshold considerable overlap is evident between returned results from distribution 121 (positive) and distribution 122 (strong positive), and thus the issue of lack of reproducibility between different methods and laboratories is not overcome. Clinicians would not have the benefit of pooled results from numerous laboratories and would be unable to differentiate between moderately high and very high TTG2 concentrations.

Measurement uncertainty is a parameter that describes the range of results that could be reasonably expected when a measurement is made [19]. In the diagnostic immunology laboratory, measurement uncertainty is usually expressed as standard deviation of the mean or normalized to the mean as coefficient of variation (CV). The reproducibility of the Inova Quantalite TTG2 antibody assay was acceptable and similar to other ELISAs in our hands, with a mean CV of 10.3% based on serial analysis of kit-independent sera over the study period. However, a larger number of random and systematic factors contrive to increase the dispersion of results when the same measurement is made in different laboratories by similar methodology. These include, but are not limited to: kit and reagent factors (different lot numbers in-use at different centres, different storage conditions); assay procedure factors (manual *versus* automated method, different automation systems, instrument calibration, in-house modifications to the protocol, different plate readers) and environmental factors (temperature control, sample storage, humidity). For these reasons, establishing a cut-off limit that can be reliably implemented between laboratories is an onerous task, even when the same methodology is used. This applies particularly to autoimmune serology, where the analyte is not a single monomorphic and well-defined chemical entity, but rather a set of different antibody combinations in different people competing for a substrate.

We next explored TTG2 antibody levels when normalized to ULN for the method across the same three distributions, in order to explore the applicability of normalized cut-offs both within and between selected methods. A cut-off at $10 \times \text{ULN}$ was chosen in accordance with recent guidelines [12]. Two clear findings emerge: first, the wide dispersion of quantitative results affects all of the immunoassays that were evaluated; secondly, normalization to ULN does not harmonize results between TTG2 methods. These two factors contrive to produce a wide dispersion of results between methods and centres, resulting in poor consensus regarding a cut-off at $10 \times \text{ULN}$.

The method divergence was particularly marked and interesting: notably, the Orgentec and Varelisa methods very rarely produced results in excess of $10 \times \text{ULN}$, in contrast to

the Inova and Euroimmun methods. The Phadia 250 method (from the same manufacturer as the Varelisa method) and the Aesku method lay between these extremes, albeit with some disagreement within the method groups. These findings highlight the arbitrary nature of the units of TTG2 antibody measurement. Kit-specific ranges are not a solution because this does not resolve problems arising from the wide dispersion of results between different laboratories using the same method.

Overall, our findings demonstrate that recommending a single cut-off for general use – whether based on a quantitative value or multiples of ULN – would result in considerable variation in patient outcome depending on location. The ESPGHAN guidelines include numerous additional required criteria for diagnosis, including a history of gluten-dependent symptoms, positive endomysial antibody on a separate occasion and confirmation of a high-risk haplotype [12]; these criteria have not been evaluated here, but our study suggests the need for better standardization of a key decision point (TTG2 antibody) in the pathway.

Despite these considerations, the principle of deferring 'gold standard' investigations is very well established in clinical medicine, and has clear benefits for patients, clinicians and the wider health economy. Our data add further support to the view that this principle can be applied to CD, but suggests that decision points based on fixed TTG2 antibody levels are currently problematic. A combination of clinical judgement and locally validated cut-offs may be preferred, in order to prevent errors of generalization and variation in outcome by location. Regular local audit of outcomes would be necessary to ensure consistency and efficacy. From an industry and laboratory perspective, there is clearly a strong case for improvements to TTG2 assay standardization, in order to meet the increasing requirements of CD serology users.

Acknowledgements

We are grateful to Professor Simon Murch (Professor of Paediatrics, Warwick Medical School) for helpful revisions to the manuscript. This study received no financial support.

Disclosures

A. B. is a member of The Coeliac Disease working group of The British Society for Paediatric Gastroenterology, Hepatology and Nutrition and was involved in producing UK guidance for implementation in children [13]. The other authors have no conflicts of interest to declare.

Author contributions

M. T. was the principle investigator. M. H., A. B., F. S., J. L. and E. S. produced the raw data. The data were analysed by L. B., M. K. and M. T. NEQAS data together with analytical

guidance was provided by W. E., R. R. S. and D. P. S. G. provided statistical support for study design and data analysis. The paper was written by L. B. and M. T. and reviewed by all authors.

References

- 1 Leffler DA, Schuppan D. Update on serologic testing in celiac disease. *Am J Gastroenterol* 2010; **105**:2520–4.
- 2 Richey R, Howdle P, Shaw E, Stokes T. Recognition and assessment of coeliac disease in children and adults: summary of NICE guidance. *BMJ* 2009; **338**:b1684.
- 3 Ciclitira P, Dewar D, McLaughlin S, Sanders D. The management of adults with coeliac disease. British Society of Gastroenterology Guidelines. Available at: www.bsg.org.uk/images/stories/clinical/bsg_coeliac_10.pdf (accessed 22 January 2014).
- 4 Green P, Rostami K, Marsh M. Diagnosis of coeliac disease. *Best Pract Res Clin Gastroenterol* 2005; **19**:389–400.
- 5 Donaldson M, Book L, Leiferman K, Zone J, Neuhausen S. Strongly positive tissue transglutaminase antibodies are associated with Marsh 3 histopathology in adult and pediatric celiac disease. *J Clin Gastroenterol* 2008; **42**:256–60.
- 6 Hill P, Holmes G. Coeliac disease: a biopsy is not always necessary for diagnosis. *Aliment Pharmacol Ther* 2008; **27**:572–7.
- 7 Vivas S, Ruiz de Morales JG, Riestra S *et al.* Duodenal biopsy may be avoided when high transglutaminase antibody titers are present. *World J Gastroenterol* 2009; **15**:4775–80.
- 8 Dahlbom I, Korponay-Szabo IR, Kovacs JB, Szalai Z, Maki M, Hansson T. Prediction of clinical and mucosal severity of coeliac disease and dermatitis herpetiformis by quantification of IgA/IgG serum antibodies to tissue transglutaminase. *J Pediatr Gastroenterol Nutr* 2010; **50**:140–6.
- 9 Zanini B, Magni A, Caselani F *et al.* High tissue-transglutaminase antibody level predicts small intestinal villous atrophy in adult patients at high risk of celiac disease. *Dig Liver Dis* 2011; **44**:280–5.
- 10 Mubarak A, Wolters V, Gmelig-Meyling F, Kate F, Houwen R. Tissue transglutaminase levels above 100 U/mL and celiac disease: a prospective study. *World J Gastroenterol* 2012; **18**:4399–403.
- 11 Nenna R, Tiberti C, Petrarca L *et al.* Anti-transglutaminase immunoreactivity and histological lesions of the duodenum in coeliac patients. *Int Immunol* 2013; **25**:389–94.
- 12 Husby S, Koletzko S, Korponay-Szabo IR *et al.* European Society for Pediatric Gastroenterology, Hepatology, and Nutrition guidelines for the diagnosis of coeliac disease. *J Pediatr Gastroenterol Nutr* 2012; **54**:136–60.
- 13 Murch S, Jenkins H, Auth M *et al.* Joint BSPGHAN and Coeliac UK guidelines for the diagnosis and management of coeliac disease in children. *Arch Dis Child* 2013; **98**:806–11.
- 14 Egner W, Shrimpton A, Sargur R, Patel D, Swallow K. ESPGHAN guidance on coeliac disease 2012: multiples of ULN for decision making do not harmonise assay performance across centres. *J Pediatr Gastroenterol Nutr* 2013; **55**:733–5.
- 15 Vermeersch P, Geboes K, Marien G, Hoffman I, Hiele M, Bossuyt X. Defining thresholds of antibody levels improves diagnosis of celiac disease. *Clin Gastroenterol Hepatol* 2013; **11**:398–403.
- 16 Giersiepen K, Legemann M, Stuhldreher N *et al.* Accuracy of diagnostic antibody tests for coeliac disease in children: summary of an evidence report. *J Pediatr Gastroenterol Nutr* 2012; **54**:229–41.
- 17 Klapp G, Masip E, Bolonio M *et al.* Celiac disease: the new proposed ESPGHAN diagnostic criteria do work well in a selected population. *J Pediatr Gastroenterol Nutr* 2013; **56**:251–6.
- 18 Ravelli A, Villanacci V. Tricks of the trade: how to avoid histological pitfalls in celiac disease. *Pathol Res Pract* 2012; **208**:197–202.
- 19 White G. Basics of estimating measurement uncertainty. *Clin Biochem Rev* 2008; **29**:S53.