ORIGINAL RESEARCH

Accuracy of a no-biopsy approach for the diagnosis of coeliac disease across different adult cohorts

Hugo A Penny,1 Suneil A Raju,1 Michelle S Lau,1 Lauren JS Marks,1 Elisabeth MR Baggus,1 Julio C Bai,2 Gabrio Bassotti,3 Hetty J Bontkes,4 Antonio Carroccio,5 Mihai Danciu,6 Mohammad H Derakhshan,7 Arzu Ensari,8 Azita Ganji,9 Peter H R Green,10 Matt W Johnson,11 Sauid Ishaq,1,12 Benjamin Lebwohl,10 Adam Levene,11 Roxana Maxim,13 Hamid Mohaghegh Shalmani,14 Mohammad Rostami-Nejad,14 David Rowlands,15 Irene A Spiridon,6 Amitabh Srivastava,16 Umberto Volta,17 Vincenzo Villanacci,18 Graeme Wild,1 Simon S Cross,1 Kamran Rostami,19 David S Sanders1

ABSTRACT

Objective We aimed to determine the predictive capacity and diagnostic yield of a 10-fold increase in serum IgA antitissue transglutaminase (tTG) antibody levels for detecting small intestinal injury diagnostic of coeliac disease (CD) in adult patients.

Design The study comprised three adult cohorts. Cohort 1: 740 patients assessed in the specialist CD clinic at a UK centre; cohort 2: 532 patients with low suspicion for CD referred for upper GI endoscopy at a UK centre; cohort 3: 145 patients with raised tTG titres from multiple international sites. Marsh 3 histology was used as a reference standard against which we determined the performance characteristics of an IgA tTG titre of ≥10×ULN for a diagnosis of CD.

Results Cohort 1: the sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) for IgA tTG levels of ≥10×ULN at identifying individuals with Marsh 3 lesions were 54.0%, 90.0%, 98.7% and 12.5%, respectively. Cohort 2: the sensitivity, specificity, PPV and NPV for IgA tTG levels of ≥10×ULN at identifying individuals with Marsh 3 lesions were 50.0%, 100.0%, 100.0% and 98.3%, respectively. Cohort 3: the sensitivity, specificity, PPV and NPV for IgA tTG levels of ≥10×ULN at identifying individuals with Marsh 3 lesions were 30.0%, 83.0%, 95.2% and 9.5%, respectively.

Conclusion Our results show that IgA tTG titres of ≥10×ULN have a strong predictive value at identifying adults with intestinal changes diagnostic of CD. This study supports the use of a no-biopsy approach for the diagnosis of adult CD.

INTRODUCTION

Coeliac disease (CD) is a common immune-mediated enteropathy, whereby increased immunological sensitivity to dietary gluten results in chronic small intestinal mucosal inflammation in genetically susceptible individuals.1 CD affects 1% of the global population, but despite its rising prevalence, the majority of patients remain undiagnosed.2 Currently, most patients with suspected CD...
are screened serologically for antibody positivity, in particular for IgA antibodies to tissue transglutaminase 2 (tTG) and antiendomysial IgA antibodies (EMAs).\(^1\) As these serological markers are not 100% specific for detecting intestinal lesions compatible with CD, positive coeliac serology is confirmed by duodenal biopsies demonstrating the hallmark pathological changes of mucosal remodelling, such as villous atrophy, crypt hyperplasia and intraepithelial lymphocytosis.\(^3\) However, upper GI endoscopy with biopsies is expensive, with costs relating directly to the procedure and indirectly to the processing and analysing of samples. The procedure is poorly tolerated by many patients\(^4\) and carries risks, including sedation complications and although rare in current practice, GI haemorrhage and/or perforation.\(^5\) Furthermore, variability exists in the identification of the main histological changes of CD, meaning cases can be missed despite duodenal biopsy sampling.\(^6\) Therefore, the identification of a non-invasive approach to the diagnosis of CD is of great interest.

The 2012 European Society for the Study of Paediatric Gastroenterology, Hepatology and Nutrition (ESPGHAN) guidelines for the diagnosis of CD challenged the necessity for duodenal biopsies in paediatric patients. They suggested that an IgA tTG antibody titre greater than 10 times the upper limit of normal (ULN), in combination with a positive EMA antibody test and compatible human leucocyte antigen (HLA) genotype, is sufficient to support a diagnosis of CD in symptomatic individuals.\(^8\) This eliminates the need for gastroscopy and its associated costs/risk in selected paediatric patients.\(^9\) The latest guidelines published this year remove the requirement for the presence of symptoms and HLA testing in the diagnostic pathway.\(^10\) This underscores the specificity of a serology-based or ‘no-biopsy’ approach for the diagnosis of paediatric CD.

Serology and biopsy

All patients from cohorts 1 and 2 underwent serological testing and duodenal biopsy within 6 weeks of each other at the Royal Hallamshire Hospital, UK. Patients were recruited prospectively from an endoscopy list that is for the investigation of GI complaints, as well as general and open-access referrals. Those with IgA deficiency, with a previous diagnosis of CD and on gluten restriction prior to diagnosis were excluded from analysis. To reduce the referral bias for CD, patients referred with positive coeliac serology (EMA positivity) in primary care were also excluded from the analysis. As a result, we considered this cohort as having low suspicion for CD.

Cohort 3 comprised a retrospective analysis of adult patients with raised tTG titres from multiple sites, involving 8 countries and comprising 11 laboratories. Cases were identified based on positive coeliac serology and had appropriate duodenal biopsy sampling. The eight countries included Argentina (n=11), Iran (n=10), Netherlands (n=10), Italy (n=44), Romania (n=10), Turkey (n=10), the USA (n=40) and the UK (n=10). In total, this provided assessments on 145 patients.

METHODS

Patients

The study comprised three adult (≥16 years old) cohorts; all individuals included in the study underwent clinical assessment and had IgA tTG serology testing alongside upper GI endoscopy with duodenal biopsies performed. Cohort 1 comprised a prospective analysis of consecutive adult patients assessed in the specialist CD clinic Royal Hallamshire Hospital, UK, from 2008 to 2019 who had a serum IgA tTG test and duodenal biopsy within 6 weeks of each other, no previous diagnosis of CD, naivety to a gluten-free diet (GFD) and IgA competency. As these individuals were assessed within the specialist coeliac clinic, they were considered as having high clinical suspicion for CD.
Ethical consideration
All identifiable medical information was removed, and all analyses were performed using anonymised data. See below for full details.

Statistical analysis
Data handling was conducted using spreadsheets in Microsoft Excel (2016); statistical analysis and graph construction were performed using GraphPad Prism V.7 (GraphPad Software, San Diego, California, USA). To enable comparison of tTG levels between different assays, tTG levels were normalised by expression as a multiple of the ULN—defined as the lowest value that was considered abnormal for each assay.19 Where indicated, receiver operating characteristic (ROC) curves were constructed, and area under the curve (AUC) was calculated to determine the accuracy of tTG levels at predicting villous atrophy. Analysis of variance or Kruskal-Wallis test, with appropriate post hoc analysis, was used to compare Marsh grades with tTG titres. A p value of <0.05 was considered statistically significant.

RESULTS
Cohort 1: patients assessed in the specialist CD clinic
A total of 740 patients fulfilled the criteria and were included in cohort 1. The prevalence of CD in this cohort was 93% (690/740 patients). Patient demographics and clinical details are listed in figure 1A. Altogether, 380/740 (51%) had tTG titres of ≥10×ULN. Marsh 3 lesions were detected in 375/380 (98.7%) of those with tTG titres of ≥10×ULN (figure 1B). Of the 740 patients, 360 (49%) had tTG titres of <10×ULN; of those, 315/360 (87.5%) had Marsh 3 lesions, 35/360 (9.7%) had Marsh 1–2 lesions and 10/360 (2.8%) had Marsh 0 lesions on duodenal biopsy. Based on this, the positive predictive value (PPV) of tTG values of ≥10×ULN at predicting mucosal abnormalities with villous atrophy was 98.7% (97.0%–99.4%) (figure 1C). Five patients had tTG titres of ≥10×ULN without detectable villous abnormalities (ie, Marsh<3 lesions); the demographic and clinical characteristics of these patients are shown in table 1. All five patients had symptoms of malabsorption and/or were EMA positive and all were positive for the HLA DQ2/8 genotype. All patients were started on a GFD and subsequently, tTG titres decreased.

Two different tTG assays were used to test tTG titres in cohort 1 over the study period; we next evaluated the differences between these two assays. To do this, we divided the cohort into two, based on the tTG assay used (cohort 1A, Aeskulisa Diagnostics; cohort 1B, ELiA Celikey). The demographics of individuals included in cohort 1A and 1B are detailed in figure 2A. Analysis of tTG titres showed that there were differences in the predictive power of the two tests, such that applying a 10×ULN cut-off had a PPV of 98.1% (95.7%–99.1%) for Marsh 3 lesions at predicting villous atrophy was 98.7% (97.0%–99.4%) (figure 1C). Five patients had tTG titres of ≥10×ULN without detectable villous abnormalities (ie, Marsh<3 lesions); the demographic and clinical characteristics of these patients are shown in table 1. All five patients had symptoms of malabsorption and/or were EMA positive and all were positive for the HLA DQ2/8 genotype. All patients were started on a GFD and subsequently, tTG titres decreased.

Two different tTG assays were used to test tTG titres in cohort 1 over the study period; we next evaluated the differences between these two assays. To do this, we divided the cohort into two, based on the tTG assay used (cohort 1A, Aeskulisa Diagnostics; cohort 1B, ELiA Celikey). The demographics of individuals included in cohort 1A and 1B are detailed in figure 2A. Analysis of tTG titres showed that there were differences in the predictive power of the two tests, such that applying a 10×ULN cut-off had a PPV of 98.1% (95.7%–99.1%) for Marsh 3 lesions at predicting villous atrophy was 98.7% (97.0%–99.4%) (figure 1C). Five patients had tTG titres of ≥10×ULN without detectable villous abnormalities (ie, Marsh<3 lesions); the demographic and clinical characteristics of these patients are shown in table 1. All five patients had symptoms of malabsorption and/or were EMA positive and all were positive for the HLA DQ2/8 genotype. All patients were started on a GFD and subsequently, tTG titres decreased.

### Table 1
Demographic and clinical details of patients with tTG titres≥10×ULN where villous abnormalities were not detected on duodenal biopsies

<table>
<thead>
<tr>
<th>Patient 1</th>
<th>Patient 2</th>
<th>Patient 3</th>
<th>Patient 4</th>
<th>Patient 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>34</td>
<td>42</td>
<td>43</td>
<td>33</td>
</tr>
<tr>
<td>Gender</td>
<td>Female</td>
<td>Male</td>
<td>Female</td>
<td>Male</td>
</tr>
<tr>
<td>Presenting symptoms</td>
<td>Diarrhoea with urgency</td>
<td>Diarrhoea, abdominal pain</td>
<td>Iron deficiency anaemia, diarrhoea</td>
<td>Abdominal pain, bloating, tiredness</td>
</tr>
<tr>
<td>HLA status</td>
<td>DQ2 heterozygous</td>
<td>DQ2 homozygous</td>
<td>DQ2 heterozygous</td>
<td>DQ2 homozygous</td>
</tr>
<tr>
<td>EMA test</td>
<td>Negative</td>
<td>Positive</td>
<td>Negative</td>
<td>Positive</td>
</tr>
<tr>
<td>Family history of CD?</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>TtG at diagnosis</td>
<td>300</td>
<td>300</td>
<td>211</td>
<td>292</td>
</tr>
<tr>
<td>Marsh grade at diagnosis</td>
<td>Marsh 1</td>
<td>Marsh 2</td>
<td>Marsh 2</td>
<td>Marsh 1</td>
</tr>
<tr>
<td>TTG post-GFD</td>
<td>1</td>
<td>8</td>
<td>6</td>
<td>111</td>
</tr>
</tbody>
</table>

CD, coeliac disease; EMA, endomysial antibody; GFD, gluten free diet; HLA, human leucocyte antigen; tTG, tissue transglutaminase.
Coeliac disease

A total of 778 patients were included in cohort 2. Ten patients had IgA deficiency; 91 patients were already consuming a GFD and/or had a diagnosis of CD; 1 patient had clinical features in keeping with a relapse of known small bowel Crohn’s disease; and 23 individuals had incomplete testing or missing information, and so these individuals were excluded from further analysis. Of the remaining cohort, 121 patients had positive EMA testing in primary care, and so these were excluded from further analysis.

Therefore, 532 patients were included in the analysis stage. The prevalence of CD in this cohort was 3.2% (17/532 patients). Patient demographics and clinical details are presented in figure 3A. 18/532 patients had Marsh 3 lesions identified on duodenal histology, 15 patients (2.8% of the cohort) were diagnosed with typical CD; 2 patients (0.4% of the cohort) had seronegative CD; and 1 patient (0.2% of the cohort) was diagnosed with non-coeliac seronegative villous atrophy.

The ELiA Celkey was used to measure tTG levels in all patients in cohort 2. Altogether, 9,532 (1.7%) had tTG titres of ≥10×ULN. Marsh 3 lesions were detected in 9,523 (100%) of those with tTG titres of ≥10×ULN (figure 3B). All individuals (9/9) with tTG titres of ≥10×ULN were EMA positive. Of the 532 patients included, 523 (98.3%) had tTG titres of <10×ULN; of those, 9,523 (1.7%) had Marsh 3 lesions; 33,523 (6.3%) had Marsh 1–2 lesions; and 481,523 (92.0%) had Marsh 0 lesions on duodenal histology. Thus, the PPV of tTG values of ≥10×ULN at predicting individuals with Marsh 3 lesions was 100% (figure 3C).

ROC analysis showed an AUC of 0.9179 for tTG values at predicting Marsh 3 histology (figure 4A). Based on this analysis, the optimal threshold of tTG titres was 5.9×ULN (tTG titres of >41.5 U/mL) (figure 4B). Reducing the threshold to this level increased the sensitivity of this approach from 50% to 61%. These results demonstrate that tTG titres of 10×ULN have excellent specificity at detecting Marsh 3 lesions in a cohort of predominantly symptomatic adult patients with a low suspicion/prevalence of CD. Reducing the tTG threshold increased the number of individuals who would have benefited from this strategy, without compromising the predictive value of this approach.

Cohort 3: international multicentre cohort

A total of 145 patients were included in cohort 3. The prevalence of CD in this cohort was 92% (133/145 patients). Demographic and clinical characteristics are included in figure 5A. Of the 145 patients included, 42 (29.0%) had tTG titres of ≥10×ULN, of which 40/42 (95.2%) patients had Marsh 3 lesions on duodenal biopsy (figure 5B). One hundred three of the 145 patients (71.0%) had tTG titres of <10×ULN; of those, 93/103 (90.3%) patients had Marsh 3 lesions; 6/103 (5.8%) patients had Marsh 1–2 lesions; and 4/103 (3.9%) had Marsh 0 lesions on duodenal

---

Figure 2  (A) Demographic and clinical details of the patients in cohorts 1A and 1B. (B) tTG titres against Marsh grade in patients from cohort 1A. (C) tTG titres against Marsh grade in patients from cohort 1B. Horizontal bar represents the mean; error bars denote SE of the mean. Black dotted line represents 1×ULN threshold; red dotted line represents 10×ULN threshold. Comparisons between groups were made with analysis of variance and post hoc Tukey test. (D) Performance characteristics of tTG titres of 10×ULN at predicting villous atrophy. **P<0.01, ****P<0.0001. tTG, tissue transglutaminase; ULN, upper limit of normal; PPV, positive predictive value; NPV, negative predictive value.
biopsy. Based on this, the predictive value of using a 10×ULN threshold at detecting individuals with Marsh 3 lesions was 95.2% (84.6%–98.7%) (figure 5C). Two patients that had tTG titres of ≥10×ULN without villous abnormalities both had Marsh 2 lesions on duodenal biopsy and were diagnosed with CD. Interestingly, tTG titres did not reach the 10×ULN threshold across three sites, despite 56 individuals having Marsh 3 lesions across these sites. Therefore, a uniform tTG threshold of 10×ULN may exclude many true positive cases.

DISCUSSION

Altogether this study included 1417 patients; 431/1417 (30%) patients had tTG titres of ≥10×ULN, of which 424 (98%) patients had Marsh 3 lesions on duodenal biopsy and were diagnosed with CD. Interestingly, tTG titres did not reach the 10×ULN threshold across three sites, despite 56 individuals having Marsh 3 lesions across these sites. Therefore, a uniform tTG threshold of 10×ULN may exclude many true positive cases.

Figure 3  (A) Demographic and clinical details of the patients in cohort 2. GI symptoms include abdominal pain, dyspepsia, dysphagia, nausea, vomiting, bloating and constipation; extraintestinal symptoms include ataxia, headache, dizziness, fatigue, joint pain and skin rash; malabsorptive symptoms include diarrhoea, haematocrit deficiencies with/without anaemia and weight loss. (B) tTG titres against Marsh grade in patients from cohort 2. Horizontal bar represents the mean; error bars denote SE of the mean. Black dotted line represents 1×ULN threshold; red dotted line represents 10×ULN threshold. Comparisons between groups were made with analysis of variance and post hoc Tukey test. (C) Performance characteristics of tTG titres of 10×ULN at predicting villous atrophy. ***P<0.0001. tTG, tissue transglutaminase; ULN, upper limit of normal; PPV, positive predictive value; NPV, negative predictive value.

Figure 4  (A) Receiver operating characteristic curve analysis of tTG values against Marsh 3 histology in cohort 2. (B) Performance characteristics of tTG titres×5.9 ULN at predicting Marsh 3 lesions. tTG, tissue transglutaminase; ULN, upper limit of normal; PPV, positive predictive value; NPV, negative predictive value.
populations (CD prevalence >70%).

In accordance with these previous studies, we found that a tTG titre threshold of 10×ULN accurately identified individuals with Marsh 3 lesions.

However, the pretest probability/disease prevalence within a cohort can influence the PPV of a diagnostic test. This means that as the disease prevalence increases, the more likely it is that a positive test represents a true positive case. The predictive value of the 10×ULN threshold in adults has less commonly been evaluated in more general clinical scenarios where the disease prevalence of CD is considered <10%. Sugai et al previously demonstrated that a tTG titre threshold of around 7×ULN had a 100% PPV for CD in an adult cohort with a CD prevalence of 3.3%. More recently, Fuchs et al demonstrated that a tTG titre threshold of 10×ULN was accurate at identifying individuals with Marsh 3 lesions across multiple sites/countries and multiple assays (cohort 3). The PPV of the 10×ULN tTG threshold in a more general clinical scenario have been estimated, but not studied.

Therefore, to confirm and extend these findings, we next sought to determine the predictive capacity of the 10×ULN threshold against the presence of Marsh 3 lesions on duodenal biopsy in a cohort of individuals considered as low suspicion for CD (cohort 2). Low suspicion was determined on the basis that, in those referred for upper GI endoscopy, the referral bias for CD was removed by excluding those who tested positive for CD in primary care. In addition, these data were generated in adult cohorts within secondary and tertiary care settings, and thus, these results do not apply to screening the general population for a diagnosis of CD. These are important considerations to ensure that patients receive the correct diagnosis and to prevent the inappropriate institution of a GFD, which has been raised as a concern regarding this approach by some.

One of the major challenges in implementing this diagnostic pathway is the variability and lack of standardisation between tTG assays. The inability to compare tTG results accurately between different assays increases the chances of misdiagnosis if a uniform threshold is used. In adults, there is a lack of data evaluating this in a multicentre international setting. To address this, we evaluated the 10×ULN threshold across multiple sites/countries and multiple assays (cohort 3). The PPV of the 10×ULN tTG threshold at detecting Marsh 3 lesions dropped from nearly 99% across cohorts 1 and 2 at our UK centre to 95.2% in a multicentre international setting. This raises the concern that sole reliance on this as a marker for CD will lead to a false-positive diagnosis in around 1 in 20 cases.

Across all cohorts, the diagnostic performance of this pathway is based strictly on the presence of Marsh 3 histology. The PPV in all three cohorts increased to >99%, where both Marsh 2 and Marsh 3 lesions are accepted as CD. It is widely accepted that in the setting of relevant symptoms/high baseline risk of CD, raised serological markers, HLA DQ2/8 positivity and/or response to a GFD, the diagnosis of CD can be established in the absence of villous effacement. Indeed, this situation may represent a false-negative histology rather than a false-negative serology due to a variety of reasons, including incorrect histological interpretation, which can vary in 10%–25% of cases between independent evaluations and/or patchy disease. Additionally, individuals with symptomatic CD may present with subtle villous abnormalities that fall within the scope of microscopic enteritis, which may lead to diagnostic inaccuracy.

In comparison with the datasets that have evaluated the tTG IgA cut-off at ≥10×ULN included in the latest paediatric ESPGHAN guidelines, 2/30 reported a PPV of <95%; 7/30 had a PPV between 95% and 99%; and 21/30 reported a PPV of >99% for detecting Marsh 3 lesions. In studies where Marsh
2 or Marsh 3 duodenal lesions were accepted for a diagnosis of CD, the PPV was >97% in all studies.9 Thus, based on the current study, the performance characteristics of this diagnostic pathway in adults with different risk levels for CD or in a multicentre international setting is at least comparable with the studies included within the paediatric guidelines.

The high tTG threshold recommended by the paediatric community in part is included to control for tTG assay variation.35 We found that for some tTG assays, a lower cut-off may retain diagnostic specificity but increase the sensitivity of this approach. In contrast, the 10×ULN threshold may not eliminate all false-positive cases across all assays. Therefore, a uniform threshold may hinder the clinical impact and accuracy of the no-biopsy approach in adult patients. Other studies have reported similar findings of differences in the optimal tTG thresholds and performance characteristics of different assays.14 19 28 This does not imply that one assay is inferior to the other nor that using tTG threshold values in this manner is not appropriate, rather that the threshold ULN cut-off should be specific to the tTG assay, as there is currently no way to accurately standardise tTG results across different assays.28 Therefore, local validation of assay specific thresholds and/or limiting the use of this pathway to certain tTG assays may help to optimise the clinical impact and accuracy of this approach.

The latest ESPGHAN guidelines no longer consider either the presence of symptoms or HLA typing as mandatory in this diagnostic pathway.9 HLA typing has been shown to add no diagnostic value in this pathway in children or adults, and testing is costly and not widely available.28 Therefore, the HLA status of individuals was not specifically evaluated in this study. In addition, while each cohort comprised predominantly symptomatic adults, we did not specifically evaluate symptomatology in this study. Current ESPGHAN guidelines recommend confirming tTG titres with EMA positivity in a second blood sample.9 EMA status was not evaluated in all patients in the present study. While EMA positivity has been shown to enhance the predictive capacity of the biopsy avoidance approach, EMA testing incurs a high cost and is labour intensive, and the interpretation is subjective.35 For these reasons, some centres have stopped performing EMA tests.17 Therefore, including EMA testing in this algorithm may lead to poor uptake of this pathway in some areas. Alternative strategies have been suggested, including simultaneous tTG measurements28 or testing for deamidated gliadin peptides (DGPs) alongside tTG titres, the latter of which has shown promise in both adult and paediatric cohorts.11 39 Further adult studies are awaited to address the place for DGPs in this algorithm.

Finally, it is noteworthy that the applicability of the biopsy avoidance approach is not as straightforward in adults as it is in children. Indeed, it should be stressed there are a number of reasons that a biopsy should still be performed in adults, particularly if there are ‘red flag’ signs/symptoms, such as, but not limited to, persistent dyspepsia in those aged ≥55 years, weight loss, severe malnutrition, dysphagia and unexplained iron deficiency anaemia.40–42 Some of these presentations align with those of adult CD; this underscores the importance of assessment by a gastroenterologist before a diagnosis of CD is made in the absence of a duodenal biopsy. In addition, a biopsy should also be considered in cases of discordance between serology and risk, and in borderline cases.42 Moreover, individuals with persistent symptoms despite adherence to a GFD require duodenal biopsy sampling, as tTG titres provide no information regarding histological improvement after introduction of a GFD.43

In conclusion, our data show that anti-tTG titres of ≥10×ULN perform well at identifying adults with intestinal changes associated with CD. This study supports a no-biopsy strategy in adult gastroenterology services. This approach has implications in reducing the cost, risk and caseload associated with diagnostic endoscopy in adult CD. However, local validation of assay-specific thresholds would enable this approach to have the greatest impact on adult patients.

Author affiliations
1Academic Unit of Gastroenterology, Royal Hallamshire Hospital, Sheffield, UK
2Medicine, Gastroenterology Hospital ‘Dr C Bonorino Udaondo’, Buenos Aires, Argentina
3Gastroenterology & Hepatology Section, Department of Medicine, University of Perugia Medical School, Perugia, Italy
4Department Clinical Chemistry, Amsterdam Gastroenterology and Metabolism and Infection and Immunity Institutes, Amsterdam UMC, Amsterdam, The Netherlands
5Department of Health Promotion Sciences, Maternal and Infant Care, Internal Medicine and Medical Specialties (PROMISE), University of Palermo, Palermo, Italy
6Pathology Department, Grigore T. Popa University Medical School, Iasi, Turkey
7Department of Gastroenterology and Hepatology, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran
8Department of Gastroenterology, Luton and Dunstable Hospital NHS Foundation Trust, Luton, UK
9Department of Gastroenterology, Dudley Group NHS Foundation Trust, Birmingham, UK
10Celiac Disease Center, Department of Medicine, Columbia University College of Physicians and Surgeons, New York, New York, USA
11Gastroenterology, Luton and Dunstable Hospital NHS Foundation Trust, Luton, UK
12Department of Gastroenterology, Ankara University Medical School, Ankara, Turkey
13Gastroenterology Department, Grigore T. Popa University of Medicine and Pharmacy Iasi, Iasi, Romania
14Gastroenterology and Liver Diseases Research Center, Research Institute for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences, Tehran, Iran
15Department of Gastroenterology, Queen Elizabeth II Hospital, Hertfordshire, UK
16Pathology, Brigham and Women’s Hospital, Boston, Massachusetts, USA
17Department of Medical and Surgical Sciences, University of Bologna, Bologna, Italy
18Department of Pathology, Spedali Civili, Brescia, Italy
19Department of Gastroenterology, MidCentral District Health Board, Palmerston North, New Zealand

Contributors HAP analysed the data and wrote the original manuscript. KR and DDS conceived the study, collected the data and edited the manuscript. SAR, ML, LISM and EMRB recruited patients and helped collect and collate the data. JB, GB, HUB, AC, MD, MHD, AE, AG, PHRG, MWI, SI, BL, AL, RM, HMS, MR-N, DR, IS, AS, UV, VV, GW and SC helped collect/generate the data. All authors approved the final version of the manuscript.

Funding BL is supported by The Louis and Gloria Fanzer Philanthropic Trust.

Competing interests DDS receives an educational grant from Dr Schär (a gluten-free food manufacturer). PHRG serves on the advisory board of ImmusanT, Cellimmune and ImmunogenX, and is an unpaid member of Nima’s Scientific Advisory Board. The remaining authors disclose no conflicts.

Patient and public involvement Patients and/or the public were not involved in the design, conduct, reporting or dissemination plans of this research.

Patient consent for publication Not required.

Ethics approval The Sheffield arm of the study (cohorts 1 and 2) was approved by the Yorkshire and the Humber Sheffield Research Ethics Committee, under registration numbers 14/YH/1216 and 15/YH/1216, respectively. Data relating to the multicentre cohort were previously approved by the research and development/audit departments of countries involved and were fully approved by the ethical committee of the Research Institute for Gastroenterology and Liver Disease, Shahid Beheshti University of Medical Science Tehran (protocol IR.SBMU.RIGLD.REC.1395.87) or the Columbia University Irving Medical Center Institutional Review Board (protocol IRB-AAA08960).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement No data are available.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is
REFERENCES