



# Validation of Antibody-Based Strategies for Diagnosis of Pediatric Celiac Disease Without Biopsy

Johannes Wolf,<sup>1,\*</sup> David Petroff,<sup>2,\*</sup> Thomas Richter,<sup>3</sup> Marcus K. H. Auth,<sup>4</sup> Holm H. Uhlig,<sup>5,6</sup> Martin W. Laass,<sup>7</sup> Peter Lauenstein,<sup>8</sup> Andreas Krahel,<sup>9</sup> Norman Händel,<sup>10</sup> Jan de Laffolie,<sup>11</sup> Almuthe C. Hauer,<sup>12</sup> Thomas Kehler,<sup>13</sup> Gunter Flemming,<sup>14</sup> Frank Schmidt,<sup>15</sup> Astor Rodrigues,<sup>6</sup> Dirk Hasenclever,<sup>16,§</sup> and Thomas Mothes<sup>1,§</sup>

<sup>1</sup>Institute of Laboratory Medicine, Clinical Chemistry and Molecular Diagnostics, Medical Faculty of the University and University Hospital, Leipzig, Germany; <sup>2</sup>Clinical Trial Centre, University of Leipzig, Leipzig, Germany; <sup>3</sup>Children's Hospital of the Clinical Centre Sankt Georg Leipzig, Leipzig, Germany; <sup>4</sup>Alder Hey Children's National Health Service Foundation Trust, Liverpool, United Kingdom; <sup>5</sup>Translational Gastroenterology Unit, Nuffield Department of Medicine, University of Oxford, John Radcliffe Hospital, Oxford, United Kingdom; <sup>6</sup>Department of Pediatrics, University of Oxford, Oxford, United Kingdom; <sup>7</sup>University Children's Hospital, Technical University Dresden, Dresden, Germany; <sup>8</sup>German Clinic for Diagnostics, Helios Children's Hospital, Wiesbaden, Germany; <sup>9</sup>Children's Hospital Prinzessin Margaret, Darmstadt, Germany; <sup>10</sup>University Children's Hospital Leipzig, Germany; <sup>11</sup>Children's Hospital, Justus Liebig University Giessen, Germany; <sup>12</sup>University Children's Hospital Graz, Austria; <sup>13</sup>Department of Pediatrics, Helios Hospital, Plauen, Germany; <sup>14</sup>Medical School, Hannover, Germany; <sup>15</sup>University Children's Hospital Halle, Halle, Germany; and <sup>16</sup>Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig, Leipzig, Germany

**BACKGROUND & AIMS:** A diagnosis of celiac disease is made based on clinical, genetic, serologic, and duodenal morphology features. Recent pediatric guidelines, based largely on retrospective data, propose omitting biopsy analysis for patients with concentrations of IgA against tissue transglutaminase (IgA-TTG) >10-fold the upper limit of normal (ULN) and if further criteria are met. A retrospective study concluded that measurements of IgA-TTG and total IgA, or IgA-TTG and IgG against deamidated gliadin (IgG-DGL) could identify patients with and without celiac disease. Patients were assigned to categories of no celiac disease, celiac disease, or biopsy required, based entirely on antibody assays. We aimed to validate the positive and negative predictive values (PPV and NPV) of these diagnostic procedures. **METHODS:** We performed a prospective study of 898 children undergoing duodenal biopsy analysis to confirm or rule out celiac disease at 13 centers in Europe. We compared findings from serologic analysis with findings from biopsy analyses, follow-up data, and diagnoses made by the pediatric gastroenterologists (celiac disease, no celiac disease, or no final diagnosis). Assays to measure IgA-TTG, IgG-DGL, and endomysium antibodies were performed by blinded researchers, and tissue sections were analyzed by local and blinded reference pathologists. We validated 2 procedures for diagnosis: total-IgA and IgA-TTG (the TTG-IgA procedure), as well as IgG-DGL with IgA-TTG (TTG-DGL procedure). Patients were assigned to categories of no celiac disease if all assays found antibody concentrations <1-fold the ULN, or celiac disease if at least 1 assay measured antibody concentrations >10-fold the ULN. All other cases were considered to require biopsy analysis. ULN values were calculated using the cutoff levels suggested by the test kit manufacturers. HLA typing was performed for 449 participants. We used models that considered how specificity values change with prevalence to extrapolate the PPV and NPV to populations with lower prevalence of celiac disease. **RESULTS:** Of the participants, 592 were found to have celiac disease, 345 were found not to have celiac disease, and 24 had no final diagnosis. The TTG-IgA procedure identified patients with celiac disease

with a PPV of 0.988 and an NPV of 0.934; the TTG-DGL procedure identified patients with celiac disease with a PPV of 0.988 and an NPV of 0.958. Based on our extrapolation model, we estimated that the PPV and NPV would remain >0.95 even at a disease prevalence as low as 4%. Tests for endomysium antibodies and HLA type did not increase the PPV of samples with levels of IgA-TTG ≥10-fold the ULN. Notably, 4.2% of pathologists disagreed in their analyses of duodenal morphology—a rate comparable to the error rate for serologic assays. **CONCLUSIONS:** In a prospective study, we validated the TTG-IgA procedure and the TTG-DGL procedure in identification of pediatric patients with or without celiac disease, without biopsy. German Clinical Trials Registry no.: DRKS00003854.

**Keywords:** AbCD Study; Endoscopy; ELISA; Gluten.

In celiac disease, gluten peptides from wheat and related cereals induce an immune-mediated enteropathy affecting 0.5%–2.5% of the population.<sup>1</sup> Celiac disease may present with a large variety of nonspecific signs and

\*Authors share co-first authorship; §Authors share co-senior authorship.

**Abbreviations used in this paper:** AbCD, Antibody Diagnostics in Paediatric Coeliac Disease trial; CI, confidence interval; EMA, endomysium antibody; ESPGHAN, European Society for Paediatric Gastroenterology, Hepatology, and Nutrition; GFD, gluten-free diet; IEL, intraepithelial lymphocytes; IgA-TTG, IgA against tissue transglutaminase; IgG-DGL, IgG against deamidated gliadin; LCB, lower 95% confidence bound; NPV, negative predictive value; PPV, positive predictive value; sIgAD, selective IgA deficiency; ULN, upper limit of normal.

Most current article

© 2017 by the AGA Institute. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

0016-5085

<http://dx.doi.org/10.1053/j.gastro.2017.04.023>

## EDITOR'S NOTES

## BACKGROUND AND CONTEXT

Guidelines permit diagnosis of Celiac disease (CD) without biopsies in children if antibody concentrations are sufficiently high ( $> 10$  upper limit of normal [ULN]), but prospective evaluation is lacking.

## NEW FINDINGS

Antibody tests alone are reliable in diagnosing (IgA-TTG  $> 10$  ULN) or excluding (IgA-TTG  $< 1$  ULN and IgG-DGL  $< 1$  ULN) CD. Antibody-based diagnoses are safe over a range of disease prevalence from about 4-60%.

## LIMITATIONS

Most study patients had prior antibody tests. Extrapolation to unselected, low prevalence cohorts is model-based.

## IMPACT

Duodenal biopsies can be avoided in a large proportion of pediatric patients with suspected CD, thereby reducing patient risks, costs, endoscopy waiting times, and delay of treatment.

symptoms. It is important to diagnose celiac disease not only in children with obvious gastrointestinal symptoms, but also in children with a less clear clinical picture because of negative health consequences.<sup>2</sup> If untreated, celiac disease is associated with reduced life expectancy.<sup>3,4</sup>

Antibody assays have been used in the diagnostics of celiac disease for nearly 60 years and although sensitivity and specificity have improved steadily,<sup>5</sup> their main role still consists in the selection of patients for subsequent diagnostic endoscopy and evaluation of intestinal histology—or in ruling out celiac disease. Such tests measure IgA against tissue transglutaminase (IgA-TTG) or against endomysium antibody (IgA-EMA) in blood. The usefulness of IgG against deamidated gliadin (IgG-DGL) peptides is still under discussion.<sup>6</sup>

Whether antibody tests can replace intestinal biopsies as the gold standard is a long-standing debate.<sup>7</sup> In 2012, the European Society for Paediatric Gastroenterology, Hepatology, and Nutrition (ESPGHAN) published new guidelines that permitted diagnosis of celiac disease without biopsies in symptomatic patients with very high IgA-TTG titers upon confirmation with IgA-EMA test and HLA typing.<sup>2</sup> The ESPGHAN suggested concentrations  $>10\times$  the upper limit of normal (ULN) for this purpose and proposed that the performance of these guidelines be evaluated prospectively.

Until now, prospective trials for evaluation of antibody tests were performed mainly with adults. Only 2 pediatric prospective studies, with 97 and 183 children, have been published. However, in the first of these studies,<sup>8</sup> all patients had celiac disease, so that predictive values could not be calculated. In the second study,<sup>9</sup> the  $10\times$  ULN rule did not lead to any false-positive findings.

Recently, we analyzed antibody data from children retrospectively and proposed 2 diagnostic procedures that resulted in high positive predictive values (PPVs), as well as negative predictive values (NPVs).<sup>10</sup> Here, we present the

results of the prospective, multicenter validation trial with 898 children.

## Methods

### Trial Registration and Ethics

The trial was registered in the German Clinical Trials Register (DRKS00003854). The ethics committee at the University of Leipzig and at each participating center approved the study protocol, patient information sheets, and informed-consent forms. The protocol and the statistical analysis plan are available in the [Supplementary Material](#), and no substantial amendments to the protocol were made. All authors had access to the study data and reviewed and approved the final manuscript.

### Sample Size

A total of 900 patients was calculated to yield a confidence interval (CI) for the PPV with a width of about 6%. This was based on an assumed prevalence of about 40% in patients fulfilling our trial's inclusion criteria.

### Study Design and Participants

AbCD (Antibody Diagnostics in Paediatric Coeliac Disease) is a prospective, multinational, multicenter, phase 3 diagnostic trial investigating patients between 5 months and 18 years scheduled for duodenal biopsy with primary aim to confirm or refute celiac disease. These inclusion criteria were intentionally chosen to be broad enough also to recruit patients who do not have celiac disease. In particular, we did not stipulate any inclusion/exclusion criteria pertaining to prior antibody tests. Patients were excluded if they already had been diagnosed with celiac disease, were on a gluten-free diet (GFD), had received immunosuppressive therapy within the last 8 weeks, were expected to be noncompliant or were participating in interventional trials. All eligible patients were asked to participate in the trial.

After documenting medical (symptoms and signs suggestive for celiac disease, local antibody test results, HLA) and family history (visit 1), blood samples were drawn and a duodenal biopsy performed (visit 2). A clinical decision regarding GFD was made without knowledge of study data, after which study data were sent to the sites, as available. After about 3 months, a follow-up took place (visit 3) comprising physical examination, documentation of clinical symptoms, dietary compliance (if on GFD), and local antibody results (if available). If blood was taken routinely to assess the effects of the GFD, an additional serum sample was collected for the study. The physician's diagnosis was documented. A further follow-up 3 months later (visit 4) took place if the physician was unable to make a diagnosis at visit 3. The study design is depicted in [Supplementary Figure 1](#).

### Serology

Sera were stored at  $-20^{\circ}\text{C}$  at the trial sites and transferred to the central laboratory (Dassow, Germany) on dry ice for blinded measurement of IgA-TTG, IgG-DGL, and IgA-EMA with test kits from EUROIMMUN (Lübeck, Germany; cutoffs  $\geq 20$  U/mL,  $\geq 25$  U/mL,  $\geq 1:10$ , respectively). IgA-EMA were

assessed by indirect immunofluorescence analysis using primate esophagus. The enzyme-linked immunosorbent assay and EMA tests were manufactured according to the regulations for in vitro diagnostic medical device after quality controls and ring trials. Duplicate measurements were performed in 2 individual dilutions of a blood sample and the results were averaged, see [Supplementary Material](#) for details about quality control. If antibody concentrations were above the measurement range, sera were serially diluted and values corrected by the dilution factor. Total IgA was assayed immunonephelometrically and concentrations  $<0.07$  g/L define selective IgA deficiency (sIgAD).<sup>11</sup>

## Histology

The study protocol recommended that the endoscopy and histology assessment should follow recent ESPGHAN guidelines,<sup>2</sup> which includes analysis of at least 1 tissue specimen from the duodenal bulb and 4 specimens from the descending duodenum. Biopsies were prepared and stained with H&E according to standard procedures at the trial sites (immunostaining by anti-CD3 or CD8 antibodies was performed routinely in 9 of 13 hospitals).

All histologic slides were first examined by a local consultant pathologist. The Marsh-Oberhuber classification was used,<sup>12</sup> but mucosal architecture and intraepithelial lymphocytes (IEL) counts were documented separately for both bulb and descending duodenum. In accordance with current guidelines,  $>25$  IEL per 100 enterocytes defined elevated IEL.<sup>2</sup> If the classification between bulb and descending duodenum differed, then the more severe histology result with the highest degree of pathology was chosen for the final result. A result with IEL elevation was always considered more severe than one without.

After the local evaluation, an independent consultant pathologist performed a second blinded evaluation of all biopsies. If results of local and study pathologist were discrepant, a third evaluation was performed by an independent pathologist. A consensus was reached upon re-evaluation if the second and third evaluations were discrepant (statistical analysis plan in supplement). The last evaluation (second, third, or consensus evaluation) was defined as final histology. If stained tissue sections did not allow for assessment due to insufficient biopsy size, numbers of biopsies, or orientation, additional tissue sections were cut and examined.

## HLA Typing

EDTA blood was collected if the patients and/or their parents consented to genetic examination. Genomic DNA of 449 participants was isolated by QIAamp DSP DNA Blood Mini Kit (Qiagen, Hilden, Germany). HLA-DQ2.2, HLA-DQ2.5, and HLA-DQ8 were determined applying the EUROArray HLA-DQ2/DQ8 kit (EUROIMMUN). Additionally, local HLA-data of 125 patients were available.

## Test Procedures

We defined 2 antibody procedures in a previous study that we are validating in the current trial.<sup>10</sup> The first requires total-IgA and IgA-TTG (TTG-IgA procedure) and the second combines IgG-DGL with IgA-TTG (TTG-DGL procedure). We use the manufacturer's cutoff ( $<1 \times$  ULN) to define clearly negative cases and define a

strictly positive test result using  $10 \times$  ULN. Values between comprise the gray zone ( $1-10 \times$  ULN). The latter needs a biopsy for confirmation or exclusion of celiac disease.

TTG-IgA procedure (sIgAD excluded): celiac disease if IgA-TTG  $\geq 10 \times$  ULN; no celiac disease if IgA-TTG  $< 1 \times$  ULN; otherwise, biopsy necessary.

TTG-DGL procedure: celiac disease if IgA-TTG  $\geq 10$  ULN or IgG-DGL  $\geq 10 \times$  ULN; no celiac disease if IgA-TTG  $< 1$  ULN and IgG-DGL  $< 1 \times$  ULN; otherwise, biopsy necessary.

## Statistics

The full analysis set includes all patients with biopsy, blood sample, and adequate follow-up. Patients without final diagnosis are not excluded. For calculation of PPV and NPV, patients without final diagnosis are considered false positive or false negative, respectively. Estimates of sensitivity or specificity are not affected. Describing characteristics of diagnostic procedures with 3 possible outcomes requires special attention ([Supplementary Material](#), statistical details). Sensitivity and specificity estimates are accompanied by 2-sided 95% CIs. Estimates of PPV and NPV are given with 1-sided 95% lower confidence bounds (LCB).

We defined a diagnostic procedure to be sufficiently reliable over a range of prevalence if PPV and NPV estimates both lie  $>95\%$  and their LCB  $>90\%$ .<sup>10</sup> The AbCD trial determined the prevalence range of reliability of the 2 procedures.

Extrapolation of PPV (NPV) to lower (higher) than observed prevalence usually uses Bayes' theorem assuming that sensitivity and specificity are independent of prevalence. Patients without celiac disease in our study population are enriched for abnormal antibody patterns leading to reduced specificity compared to an unselected population.<sup>13,14</sup> This dependence on prevalence is taken into account here by allowing for specificity in Bayes' theorem approach to depend on prevalence. In particular, we choose specificity to be a linear function of prevalence with the observed value at the observed prevalence and with a value at low prevalence taken from data of about 17,000 patients<sup>15</sup> and corroborated using a model-based analysis of the AbCD data, see statistical details in the [Supplementary Material](#).

CIs of proportions are determined using a Wilson score<sup>16</sup> and LCBs for PPV and NPV use a logit transformation.<sup>17</sup> Data were analyzed using R software, version 3.3.1.<sup>18</sup>

## Results

### Study Participants

A total of 949 children and adolescents (aged 10 months to 18 years) were enrolled from October 2012 to December 2015 by 13 children's hospitals from 3 countries. The full analysis set comprises 898 patients (529 celiac disease, 345 patients without celiac disease, and 24 without final diagnosis; [Supplementary Figure 2](#)). In almost all the patients, there were fewer than 4 days' delay between blood sample and biopsy (859 of 898 [95.7%]) and only 4 patients had blood samples taken outside the time window 2 months before the biopsy to 2 weeks afterward. [Supplementary Figure 3](#) details accrual and diagnosis by center.

Baseline characteristics are presented in [Table 1](#). The proportion of patients with diabetes mellitus type 1, Down

**Table 1.** Baseline Demographic and Clinical Characteristics

Characteristic	CD patients (n = 529)	Non-CD patients (n = 345)	No final diagnosis (n = 24)
Girls, n (%)	347 (66)	194 (56)	9 (38)
Age, y, mean (SD)	8.6 (4.5)	10.2 (5.0)	6.8 (4.6)
BMI-SDS, <sup>a</sup> mean (SD)	-0.28 (1.16)	-0.20 (1.26)	-0.35 (1.08)
<10th percentile, n (%)	88 (18)	55 (16)	6 (25)
<3rd percentile, n (%)	44 (9)	24 (7)	0 (0)
Height-SDS, <sup>a</sup> mean (SD)	-0.54 (1.29)	-0.43 (1.35)	-0.42 (1.07)
<10th percentile, n (%)	128 (26)	84 (25)	5 (21)
<3rd percentile, n (%)	71 (14)	51 (15)	2 (8)
First-degree relative with CD, n (%)	68 (13)	22 (6)	2 (8)
Diabetes mellitus type 1, n (%)	44 (8)	11 (3)	3 (12)
Down syndrome, n (%)	8 (1.5)	0 (0)	0 (0)
Autoimmune thyroid disease, n (%)	10 (1.9)	6 (1.7)	0 (0)
Turner syndrome, n (%)	3 (0.6)	0 (0)	0 (0)
Autoimmune liver disease, n (%)	1 (0.2)	1 (0.3)	0 (0)
Selective IgA-deficiency, n (%)	6 (1.1)	10 (2.9)	0 (0)
No symptoms, n (%)	59 (11)	14 (4)	1 (4)

BMI, body mass index; CD, celiac disease; IgA, immunoglobulin A; SDS, standard deviation score based on age and sex.

<sup>a</sup>BMI-SDS and Height-SDS were calculated from German reference data.<sup>19,20</sup> Significance tests were not performed because they are difficult to interpret in a heavily selected study population. See [Supplementary Table 1](#) for symptoms and signs suggestive of CD.

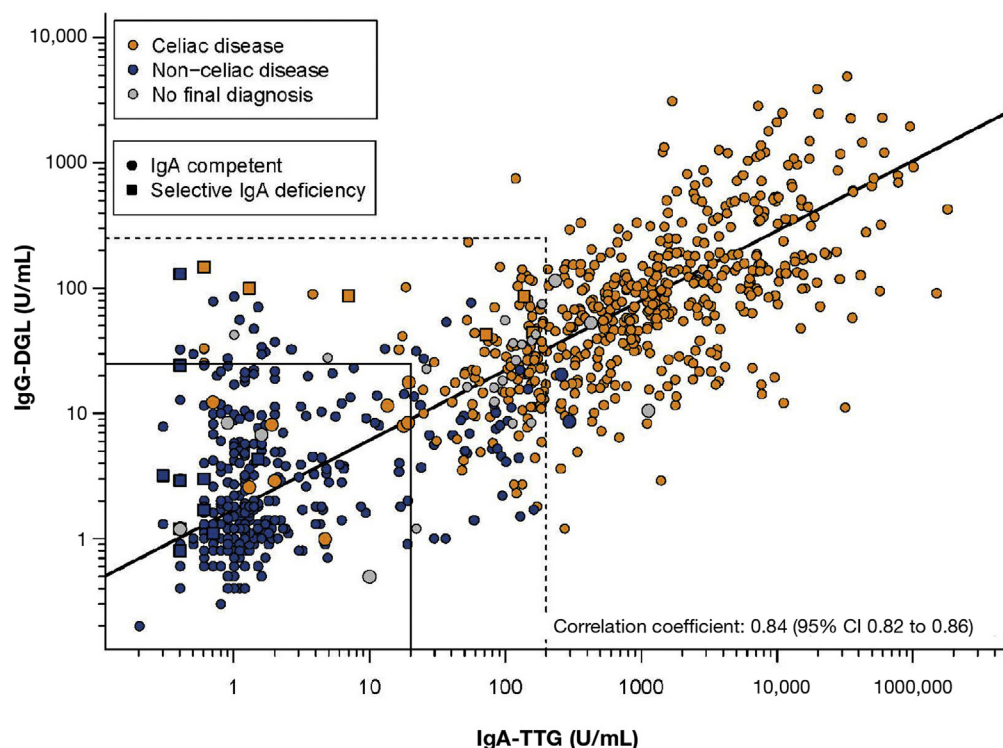
syndrome, and Turner syndrome is higher in celiac disease than in nonceliac disease cases. Signs and symptoms suggestive for celiac disease<sup>2</sup> were documented for all patients ([Supplementary Table 1](#)). Abdominal pain was the most frequent symptom in the whole study population and in patients with celiac disease as well. Symptoms provide almost no indication as to whether or not the patient has celiac disease in this heavily selected population. Iron deficiency anemia was the only reason for inclusion, which was

found more often in patients with celiac disease than in those without.

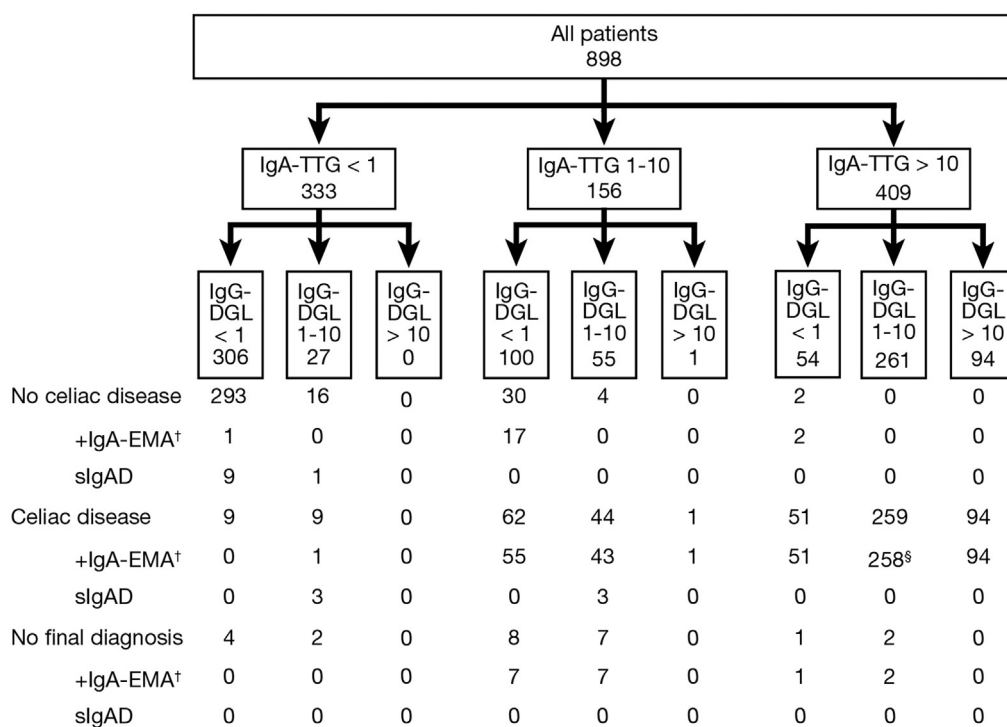
### Antibody Results at Time of Endoscopy

[Figure 1](#) depicts serum antibody concentrations and diagnoses at the time of endoscopy also indicating sIgAD patients. Note that concentrations exceeded 10,000 times the normal level in many celiac disease patients. The

**Figure 1.** Scatterplot showing the relation between IgA-TTG and IgG-DGL. The solid and dashed rectangles denote 1× ULN and 10× ULN, respectively. A linear correlation coefficient was calculated and the corresponding CI was determined using the Fisher transformation. An orthogonal regression line was fitted and the correlation analyses were performed after excluding sIgAD patients.







classification relative to the cuts used in the procedures can be found in Figure 2, where the results of IgA-EMA are also provided. The vast majority of non-celiac disease patients were  $<1 \times$  ULN for both IgA-TTG and IgG-DGL (293 of 345 children [84.9%]), and most celiac disease-patients were  $\geq 10 \times$  ULN for IgA-TTG (404 of 529 children [76.4%]).

Table 2 shows the diagnostic characteristics of the tests at the manufacturer's cutoffs, which can be calculated from the information in Figure 2.

The predictive values of the 2 test procedures to be validated can be found in Table 3. The TTG-DGL procedure meets our reliability criteria in that both PPV and NPV estimates lie  $>95\%$  and their LCB  $>90\%$ .<sup>10</sup> The PPV of both procedures are essentially identical (404 of 409 for the TTG-IgA procedure, 405 of 410 for the TTG-DGL procedure). As described in the Methods section, high predictive values are found although patients without final diagnosis, but

classified by the procedures to be positive or negative (ie, not in the gray zone), are taken to be false positive or negative according to the definitions of PPV/NPV. Less than a quarter of patients show intermediate antibody values (between  $1 \times$  ULN and  $10 \times$  ULN) and thus require a biopsy. The NPV of the TTG and IgA-procedure at the observed prevalence of 59% does not quite meet our reliability criteria.

Local antibody data from a blood sample drawn no more than a week before the one for the trial were available from 326 patients. The results of tests from various manufacturers (a list of test kits is presented in Supplementary Table 2) were pooled after converting to multiples of ULN and show that PPV is similar to that of the reference (0.984 TTG-IgA procedure, 0.963 TTG-DGL procedure), though the proportion of patients with concentrations  $>10 \times$  ULN is somewhat lower (0.77 based on reference tests vs

**Table 2.** Performance Assays of IgA against Tissue Transglutaminase, IgG against Deamidated Gliadin Peptides, and IgA Endomysium Antibodies at Manufacturer's Cutoff

Variable	IgA-TTG <sup>a</sup>	IgG-DGL <sup>b</sup>	IgA-EMA <sup>a</sup>
Sensitivity estimate (95% CI)	0.971 (0.953–0.983)	0.769 (0.732–0.803)	0.958 (0.937–0.972)
Specificity estimate (95% CI)	0.893 (0.855–0.921)	0.942 (0.912–0.962)	0.940 (0.910–0.961)
PPV, estimate (LCB)	0.904 (0.881)	0.929 (0.906)	0.931 (0.911)
NPV, estimate (LCB)	0.934 (0.908)	0.707 (0.670)	0.916 (0.888)
PPV, <sup>c</sup> estimate (LCB)	0.934 (0.914)	0.953 (0.933)	0.962 (0.945)
NPV, <sup>c</sup> estimate (LCB)	0.952 (0.928)	0.727 (0.691)	0.935 (0.909)

<sup>a</sup>Children with sIgAD were excluded before analysis.

<sup>b</sup>Children with sIgAD were not excluded before analysis.

<sup>c</sup>For comparison purposes with other publications, PPV and NPV were also determined for the subset of patients with a final diagnosis. The restriction to this subset has no impact on the estimates for sensitivity or specificity.

**Table 3.** Predictive Values of the Procedures to Be Validated

Variable	TTG-IgA procedure <sup>a</sup>	TTG-DGL procedure <sup>b</sup>
PPV, estimate (LCB)	0.988 (0.975)	0.988 (0.975)
NPV, estimate (LCB)	0.934 (0.908)	0.958 (0.934)
PPV, <sup>c</sup> estimate (LCB)	0.995 (0.984)	0.995 (0.984)
NPV, <sup>c</sup> estimate (LCB)	0.952 (0.929)	0.970 (0.950)
Proportion of patients to be biopsied, <sup>d</sup> estimate (95% CI)	0.173 (0.150–0.200)	0.203 (0.178–0.230)

<sup>a</sup>Children with sIgAD were excluded before analysis.  
<sup>b</sup>Children with sIgAD were not excluded before analysis.  
<sup>c</sup>For comparison purposes with other publications, PPV and NPV were also determined for the subset of patients with a final diagnosis.  
<sup>d</sup>Patients to be biopsied are children with antibody values in the gray zone (1× ULN to 10× ULN). A comparison with estimates of other test manufacturers is presented in [Supplementary Table 3](#).

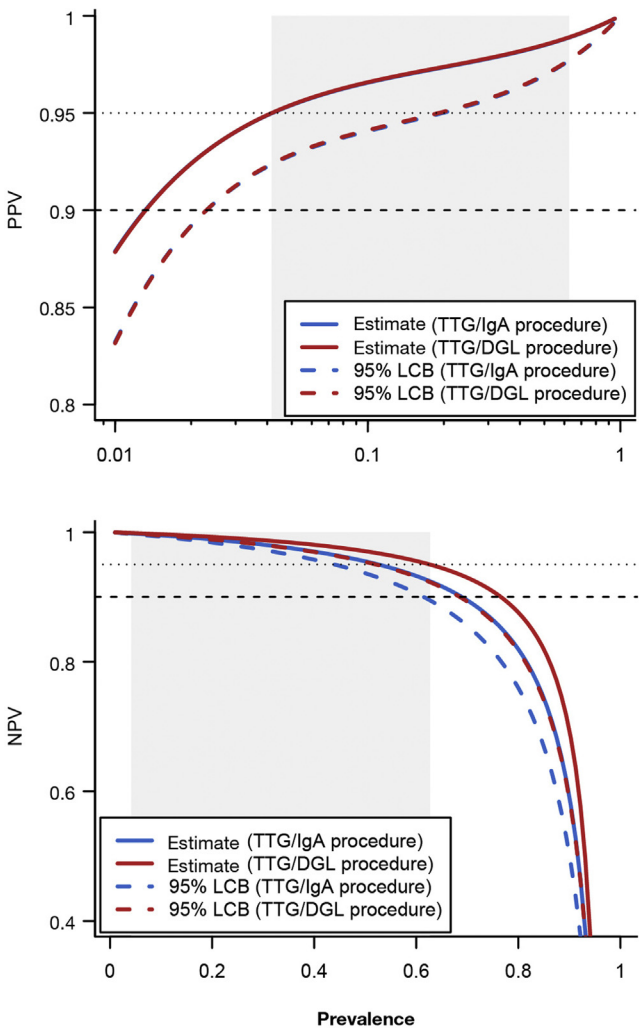
0.62–0.67 based on local tests), see [Supplementary Table 3](#). This results in a higher proportion of patients to be biopsied. NPV of the local tests may be lower than that of the reference (0.899 TTG-IgA procedure, 0.948 TTG-DGL procedure).  
We are interested in understanding the properties of the test procedures for populations that could differ substantially from ours in terms of prevalence, for example, before antibody testing, where prevalence may be <10%. PPV and NPV, as they depend on prevalence of celiac disease, are provided in [Figure 3](#) using the extrapolation technique described in the [Supplementary Material](#). The procedures are reliable according to our preset criteria for prevalence from 0.04 to 0.53 and from 0.04 to 0.63 for the TTG-IgA and TTG-DGL procedure, respectively.

**Histologic Evaluation**

In 38 of 898 cases (4.2%), there are discrepancies between the local and the first reference pathologist’s evaluations ([Table 4](#)) that are relevant for the diagnosis. In 14 cases, discrepant reading involves Marsh 1 in one assessment and Marsh 2 or 3 in the other. The discrepancy rate between local and final evaluations is 3.2%. In 18 (2%) cases, the final trial histologic evaluation conflicts with the final diagnosis (5 celiac disease, 13 not celiac disease; [Supplementary Table 4](#)).  
The agreement between the first blinded reference histologic evaluation and the results of the antibody procedures is roughly comparable with the agreement between the latter and the final diagnosis: there are 6 celiac disease patients according to the antibody procedures without histologic signs of celiac disease. Seventeen patients are categorized as non-celiac disease according to the TTG-DGL procedure, but are compatible with Marsh 3 ([Supplementary Table 5](#)).

**Discrepant Cases and Those Without Final Diagnosis**

There are only 2 patients classified as false-positive (one with type 1 diabetes mellitus) and 3 classified as positive without final diagnosis using the TTG-IgA and TTG-DGL



**Figure 3.** Predictive values of the test procedures as they depend on prevalence. PPV and NPV (solid lines) plotted as functions of the prevalence together with a 95% LCB (dashed lines). In the upper panel the procedures are virtually indistinguishable for PPV (404 of 409 in the TTG-IgA procedure and 405 of 410 in the TTG-DGL procedure at the observed prevalence of 59%, blue curve covered by red curve). The range of disease prevalence for which the TTG-DGL procedure is reliable (4%–63%) is shown by gray shading.

**Table 4.** Comparison of Results of Local and First Reference Histologic Evaluation

Local evaluation	First re-evaluation							Total
	No assessment possible <sup>a</sup>	Normal	Not suggesting celiac disease <sup>b</sup>	Marsh 1	Marsh 2	Marsh 3A	Marsh 3B and 3C	
No assessment possible <sup>a</sup>	<b>0</b>	1	0	0	0	1	0	2
Normal	1	<b>266</b>	20	9	1	9	2	308
Not suggesting celiac disease <sup>b</sup>	0	1	<b>1</b>	0	0	0	1	3
Marsh 1	0	28	1	<b>10</b>	3	6	1	49
Marsh 2	0	0	0	1	<b>2</b>	0	2	5
Marsh 3A	1	5	1	3	8	<b>47</b>	35	100
Marsh 3B and 3C	0	1	2	0	5	65	<b>358</b>	431
Total	2	302	25	23	19	128	399	<b>898</b>

NOTE. Discrepancies that may be relevant for the diagnosis are in boldface type. Based on whether or not there were such discrepancies, Cohen's measure of inter-rater agreement was  $\kappa = .912$ ; 95% CI, 0.884–0.939.

<sup>a</sup>Insufficient quality of inspected tissue, no information on mucosal architecture or IEL number.

<sup>b</sup>Crypt hyperplasia to total villous atrophy without increase of IEL.

procedures (Supplementary Table 6). Local IgA-TTG is positive ( $>2 \times \text{ULN}$ ) for all these patients, confirmed by centrally measured IgA-EMA, but histologic evaluations do not clearly suggest celiac disease, despite availability of an adequate number of biopsies.

In 21 patients (15 celiac disease and 6 without final diagnosis) the TTG-IgA procedure produces false-negative results (Supplementary Table 7), 6 of whom have had positive IgA-TTG results when performed locally. None of these patients has sIgAD. In contrast to the TTG-IgA procedure, only 13 (9 with celiac disease) are found to be negative using the TTG-DGL procedure. Twenty of the 21 patients were symptomatic and all celiac disease patients showed remission under GFD. In 2 of the false-negative celiac disease patients, there are major discrepancies between histologic evaluations. Two of 8 celiac disease patients with Down syndrome were negative for IgA-TTG (1 with IgG-DGL  $>1 \times \text{ULN}$ ) despite IgA competency.

The treating physicians were unable to make a final diagnosis in 24 patients (Supplementary Table 8) within the time frame of the study. Three of these patients would be classified as celiac disease according to the TTG-IgA and TTG-DGL procedures. In all 3, HLA-type is compatible with celiac disease and the EMA test was positive (at least 1:100). Two of the 3 cases have a final histology compatible with celiac disease. Six patients would have been classified as non-celiac disease according to the TTG-IgA procedure, though only 4 using the TTG-DGL procedure.

### *IgA–Endomysium Antibody and HLA Typing*

Of the 409 patients in our trial with IgA-TTG  $\geq 10 \times \text{ULN}$ , IgA-EMA was positive in all cases (IgA-EMA unavailable in 1 patient) and HLA status was compatible with celiac disease in all 277 cases for which it was available.

### *Asymptomatic Patients*

Seventy-four patients in our study were asymptomatic (59 celiac disease, 14 patients without celiac disease, and 1

no final diagnosis). The reasons for their inclusion are detailed in Supplementary Table 9. Most of them are patients with associated disorders (mainly type 1 diabetes mellitus), first-degree relatives of celiac disease patients, or children with prior accidental finding of increased antibodies. All but 3 of the patients had prior elevated antibody tests. Regarding diagnostic performance, there was no indication that this subgroup differed in any meaningful way from symptomatic patients. Of the 47 asymptomatic patients with IgA-TTG  $\geq 10 \times \text{ULN}$ , 46 were diagnosed with celiac disease, translating to a PPV for both procedures of 0.979. The symptomatic celiac disease patients were not more likely than the asymptomatic celiac disease patients to be classified as positive according to the test procedures (359 of 470 [76%] vs 47 of 59 [80%]).

## **Discussion**

AbCD is the first prospective trial providing evidence on the reliability of antibody tests in diagnostics of celiac disease based on a large number of children and adolescents. We show that antibody assays alone allow safe diagnosis or exclusion of celiac disease in more than three-quarters of children.

We validated 2 procedures, the TTG-IgA procedure based on assay of IgA-TTG and of total IgA, and the TTG-DGL procedure based on assay of IgA-TTG and of IgG-DGL. Both procedures are comparable and reliable regarding PPV. However, the TTG and DGL procedure is superior regarding NPV in that it can be used at higher prevalence.

A further advantage of the TTG-DGL procedure is its ability to detect IgA-TTG–negative, so-called seronegative celiac disease patients.<sup>21</sup> We found 6 IgA-competent, but IgA-TTG–negative celiac disease patients who were positive for IgG-DGL ( $>1 \times \text{ULN}$ ). Three of these patients had very low IgA-TTG ( $<0.2 \times \text{ULN}$ , see Supplementary Table 7). Such patients were also described by Dahle and colleagues<sup>22</sup> for an adult cohort. Although rare, such patients may be under-represented in our data because the majority of patients

were included based on prior positive IgA-TTG tests. Using the TTG-DGL procedure, a biopsy would be avoided in 9 of the 10 patients without celiac disease with sIgAD and all 6 celiac disease patients with sIgAD would be assigned to be biopsied.

Our trial population is not representative of a general clinical population presenting in a community pediatrician's clinic, but is strongly selected by antibody pretesting, resulting in a high celiac disease prevalence (59%). In this way, predominantly those patients with low antibody concentrations are excluded from the AbCD trial and atypical non-celiac disease patients are enriched, which leads to a lower specificity of IgA-TTG at company cutoff (Table 2, see also [Supplementary Material](#)) than usually reported.<sup>23,24</sup> This is in agreement with a recent report<sup>15</sup> using a test by the same manufacturer and with a similar selection process. Preselection also reduces the specificity of IgA-EMA, which is also lower in our study than commonly described.<sup>24</sup>

Using our 2 antibody test procedures, biopsies would have been unnecessary in three-quarters of our population. In a screening population, this number would increase to >95% because it is dominated by healthy patients, where the high specificity of the test would suffice to exclude celiac disease and avoid biopsies. Concentrating only on celiac disease patients, we found that 23% of patients in our trial would have required an endoscopy to confirm or exclude celiac disease. In contrast, in a study analyzing consecutive serologic data, a much higher proportion (57%) would have needed biopsies based on the frequency of seropositive patients with titers  $<10 \times \text{ULN}$ .<sup>15</sup>

Note, however, that avoiding a biopsy has the (low) risk of missing additional pathologies, especially in symptomatic patients that an endoscopy could have disclosed.<sup>25</sup> Therefore, a careful follow-up was recommended of symptomatic patients who had been diagnosed without a biopsy in order to avoid overlooking associated conditions.

Histology was regarded as a cornerstone of celiac disease diagnostics until the new guidelines appeared in 2012. They recommend that symptoms, celiac disease-specific antibodies, HLA, and biopsy findings should all be taken into account when diagnosing celiac disease.<sup>2</sup> Histologic evaluation of intestinal biopsies for the diagnosis of celiac disease bears a non-negligible risk of misdiagnosis.<sup>26</sup> Inter-observer discrepancy rates for diagnosis or exclusion of celiac disease are in the order of 5%–10%.<sup>26,27</sup> Our rate for major discrepancies, important for diagnosis, was 4.2% (local vs first reference evaluation). The discordance rate between either diagnostic antibody procedure and the clinical diagnosis is of the same order of magnitude as the discordance rate between pathologists. Concordance between antibody-based and histology-based diagnosis cannot be expected to reach 100% because both are error prone.

In order to forego a biopsy, the ESPGHAN guidelines recommend confirming the result IgA-TTG  $\geq 10 \times \text{ULN}$  with an IgA-EMA test and HLA genotyping. IgA-EMA was positive in all 408 AbCD patients with IgA-TTG  $\geq 10 \times \text{ULN}$ , confirming our retrospective results.<sup>10</sup> HLA status was compatible with celiac disease<sup>28</sup> in all 277 cases with IgA-TTG  $\geq 10 \times \text{ULN}$  for which this information was available.

Thus, our results indicate that confirmatory assays are unnecessary as a diagnostic feature per se. However, it is important to confirm the serologic diagnosis on an independent blood sample in order to exclude sample mix-up before starting with a lifelong GFD. To simplify the current diagnostic proposal, a positive IgA-TTG result ( $>10 \times \text{ULN}$ ) can be confirmed by a second IgA-TTG enzyme-linked immunosorbent assay test.

Of note, in case of patients with intermediate (gray zone) antibody concentrations, HLA typing still remain important to rule out false-positive results.

The guidelines also require presence of symptoms for omitting diagnostic biopsy. Seventy-four patients in our study were asymptomatic, and there was no indication that this subgroup differed in any meaningful way from the remainder, apart from having a higher prevalence of celiac disease. This may not hold for other populations, for example, those without a prior antibody test.

A limitation of the study is that the clinical diagnosis is not based exclusively on histology; knowledge of antibody results may have played a role. The agreement between the blinded final histologic assessment and the diagnosis (856 of 874 cases) is excellent, demonstrating that this problem is marginal. Upon scrutiny, antibody results could have affected the diagnosis in 2 of the 18 discrepant cases (see [Supplementary Table 4](#)). Exclusion of these cases would eliminate 1 false negative and 1 true negative with no consequences for PPV and essentially none for NPV.

A further limitation is the selection of patients described here, meaning that extrapolation was necessary to infer diagnostic properties for settings with lower prevalence. Many patients had had a positive antibody test before the trial. However, diagnostic properties of the test procedures upon examining the first blood sample are required. Thus, we extrapolated to lower prevalence based on models from the trial data and from external sources,<sup>15</sup> and such a process is uncertain to a degree. However, those considerations are not irrelevant because, according to the ESPGHAN criteria, a serologic confirmation is needed and the test criteria of our study would be in line with confirmation setting of a referral center. Sensitivity analyses were performed, however, to ensure that claims made for low prevalence err on the side of caution. This trial was designed with primarily the tests of one manufacturer in mind. The results from a fairly large number of data available from other manufacturers give reason to expect that PPV will be comparable when tests of other manufacturers are used, that is, antibody test results  $\geq 10 \times \text{ULN}$  are likely to be reliable. However, NPV may not be comparable. For confirmation, a larger number of patients has to be evaluated with different test kits.

Until now, there are only 2 pediatric, prospective studies evaluating the  $10 \times \text{ULN}$  rule with a small number of patients. However, the authors of the first prospective trial<sup>8</sup> could not assess false-positive serologic results because they included celiac disease patients only. The second study<sup>9</sup> was a monocentric trial with 120 pediatric celiac disease patients and 63 patients without celiac disease. In both papers, only clearly defined patients (celiac disease vs



non-celiac disease) are described. However, in several cases, the diagnostic process may not lead to a final decision within a reasonable time. Such cases are more likely to have contradictory antibody test results. Unclear patients are commonly neglected in diagnostic trials and it is a strength of our study not to have excluded those patients. Doing so results in artificially high PPV and NPV (Table 3).

To summarize, the prospective AbCD study demonstrates the reliability of antibody tests in diagnostics of celiac disease in children and adolescents. The TTG-DGL procedure (IgA-TTG or IgG-DGL  $\geq 10 \times$  ULN or both tests  $< 1 \times$  ULN) based on the test kits used here can reliably diagnose or exclude celiac disease. The TTG-IgA procedure (IgA-TTG  $\geq 10 \times$  ULN or  $< 1 \times$  ULN for IgA-competent patients) is almost identical in its ability to diagnose celiac disease, but has a slightly narrower prevalence range over which it can reliably exclude celiac disease. Our extrapolation shows that the 2 procedures are valid for a large range of prevalence down to 4%. Therefore, they are useful for diagnostics by the pediatric gastroenterologist without prior antibody tests. Our results have major personal and health care implications in clinical practice (avoiding many biopsies, reducing costs, endoscopy waiting times, patient risks, and delay of treatment). We have shown that HLA typing, as well as EMA tests, are not required in unequivocal cases and that endoscopic procedures to assess duodenal biopsies are not required in a substantial proportion of pediatric patients with suspected celiac disease.

## Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Gastroenterology* at [www.gastrojournal.org](http://www.gastrojournal.org), and at <http://dx.doi.org/10.1053/j.gastro.2017.04.023>.

## References

- Mustalahti K, Catassi C, Reunanen A, et al. The prevalence of celiac disease in Europe: results of a centralized, international mass screening project. *Ann Med* 2010; 42:587–595.
- Husby S, Koletzko S, Korponay-Szabó I, et al. European Society for Pediatric Gastroenterology, Hepatology, and Nutrition guidelines for the diagnosis of coeliac disease. *J Pediatr Gastroenterol Nutr* 2012;54:136–160.
- Tio M, Cox MR, Eslick GD. Meta-analysis: coeliac disease and the risk of all-cause mortality, any malignancy and lymphoid malignancy. *Aliment Pharmacol Ther* 2012; 35:540–551.
- Rubio-Tapia A, Ludvigsson JF, Choung RS, et al. Increased mortality among men aged 50 years old or above with elevated IgA anti-transglutaminase antibodies: NHANES III. *BMC Gastroenterol* 2016;16:136.
- Berger E. Cöliakie und Antikörper gegen Nahrungsmittel-Antigene. *Ann Paediatr* 1957;188:297–300.
- Brusca I. Overview of biomarkers for diagnosis and monitoring of celiac disease. *Adv Clin Chem* 2015; 68:1–55.
- Carroccio A, Iacono G, Montalto G, et al. Immunologic and absorptive tests in celiac disease: can they replace intestinal biopsies? *Scand J Gastroenterol* 1993;28: 673–676.
- 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; 15:4775–4780.
- Mubarak A, Wolters VM, Gmelig-Meyling FHJ, et al. Tissue transglutaminase levels above 100 U/mL and celiac disease: a prospective study. *World J Gastroenterol* 2012;18:4399–4403.
- Wolf J, Hasenclever D, Petroff D, et al. Antibodies in the diagnosis of coeliac disease: a biopsy-controlled, international, multicentre study of 376 children with coeliac disease and 695 controls. *PLoS One* 2014;9:e97853.
- Yel L. Selective IgA deficiency. *J Clin Immunol* 2010; 30:10–16.
- Oberhuber G, Granditsch G, Vogelsang H. The histopathology of coeliac disease: time for a standardized report scheme for pathologists. *Eur J Gastroenterol Hepatol* 1999;11:1185–1194.
- Moons KG, van Es GA, Deckers JW, et al. Limitations of sensitivity, specificity, likelihood ratio, and Bayes' theorem in assessing diagnostic probabilities: a clinical example. *Epidemiology* 1997;8:12–17.
- Leeftang MMG, Bossuyt PMM, Irwig L. Diagnostic test accuracy may vary with prevalence: implications for evidence-based diagnosis. *J Clin Epidemiol* 2009;62: 5–12.
- Gidrewicz D, Potter K, Trevenen CL, et al. Evaluation of the ESPGHAN celiac Guidelines in a North American pediatric population. *Am J Gastroenterol* 2015;110: 760–767.
- Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. *Stat Med* 1998;17:857–872.
- Mercaldo ND, Lau KF, Zhou XH. Confidence intervals for predictive values with an emphasis to case-control studies. *Stat Med* 2007;26:2170–2183.
- R Development Core Team. A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing, 2016.
- Kromeyer-Hauschild K, Wabitsch M, Kunze D, et al. Percentiles of body mass index in children and adolescents evaluated from different regional German studies. *Monatsschr Kinderheilkd* 2001;149:807–818.
- Rosario AS, Schienkiewitz A, Neuhauser H. German height references for children aged 0 to under 18 years compared to WHO and CDC growth charts. *Ann Hum Biol* 2011;38:121–130.
- Volta U, Caio G, Boschetti E, et al. Seronegative celiac disease: shedding light on an obscure clinical entity. *Dig Liver Dis* 2016;48:1018–1022.
- Dahle C, Hagman A, Ignatova S, et al. M. Antibodies against deamidated gliadin peptides identify adult coeliac disease patients negative for antibodies against endomysium and tissue transglutaminase. *Aliment Pharmacol Ther* 2010;32:254–260.

23. van Meensel B, Hiele M, Hoffman I, et al. Diagnostic accuracy of ten second-generation (human) tissue transglutaminase antibody assays in celiac disease. *Clin Chem* 2004;50:2125–2135.
24. Hill ID. What are the sensitivity and specificity of serologic tests for celiac disease? Do sensitivity and specificity vary in different populations? *Gastroenterology* 2005;128(Suppl):S25–S32.
25. Guandalini S, Newland C. Can we really skip the biopsy in diagnosing symptomatic children with celiac disease? *J Pediatr Gastroenterol Nutr* 2013;57:e24.
26. Montén C, Bjelkenkrantz K, Gudjonsdottir AH, et al. Validity of histology for the diagnosis of paediatric coeliac disease: a Swedish multicentre study. *Scand J Gastroenterol* 2016;51:427–433.
27. Arguelles-Grande C, Tennyson CA, Lewis SK, et al. Variability in small bowel histopathology reporting between different pathology practice settings: impact on the diagnosis of coeliac disease. *J Clin Pathol* 2012; 65:242–247.
28. Mubarak A, Spierings E, Wolters V, et al. Human leukocyte antigen DQ2.2 and celiac disease. *J Pediatr Gastroenterol Nutr* 2013;56:428–430.

Author names in bold designate shared co-first authorship.

Received December 7, 2016. Accepted April 19, 2017.

#### Reprint requests

Address requests for reprints to: Thomas Mothes, PhD, Institute of Laboratory Medicine, Clinical Chemistry and Molecular Diagnostics, Medical Faculty of the University and University Hospital Leipzig, Liebigstraße 27, 04103 Leipzig, Germany. e-mail: [mothes@medizin.uni-leipzig.de](mailto:mothes@medizin.uni-leipzig.de); fax: (49) 341 9722379.

#### Acknowledgments

The study was funded by the State of Mecklenburg-Vorpommern and the European Union (European Regional Development Fund) and EUROIMMUN Medizinische Labordiagnostika AG. V. Wiechmann (Department Pathology of the Clinical Centre Sankt Georg Leipzig, D. Aust, Department Pathology of the Technical University Dresden, and A. Höhn (Department Pathology of the University Hospital Leipzig, Germany), are thanked for reference pathology. The authors are very grateful to the large number of colleagues, who have played a valuable role to conduct and complete the study (J. Thiery from

Institute for Laboratory Medicine, Clinical Chemistry and Molecular Diagnostics from the University Hospital Leipzig, Y. Breitenstein and T. Haage from Coordination Centre for Clinical Trials of the University of Leipzig, Gaby Prazienka from University Hospital Leipzig, M. Zurek, E. Bertko, A. Schumann, C. Goldhardt, C. Spranger, C. Doelz, and A. Poege from Children's Hospital of the Clinical Centre Sankt Georg Leipzig, S. Leube and K. Weber in Children's Hospital of the Technical University Dresden, T. Weidenhausen, C. Wendt, R. Schwarz, Y. Kho, N. Erdem, J. Riechmann, and K.-P. Zimmer from Children's Hospital, Justus Liebig University Giessen, K.-M. Keller from Helios Children's Hospital Wiesbaden, W.-H. Huber from University Hospital Vienna, S. Thanner-Lechner, M. Hoffmann, A. Deutschmann, V. Noori-Zandi, and B. Kohlmaier from University Children's Hospital Graz, J. Chivenga from the Translational Gastroenterology Unit, Nuffield Department of Medicine Oxford and the colleagues from Biomedical Research Centre—funded Oxford Gastrointestinal Biobank for support, L. Bailey, N.M. Auth, J. McPartland, C. Tzivnikos, S.J. Allen, P. Newland, K. Padmore, and E. Jones from Alder Hey Children's National Health Service Foundation Trust in Liverpool, A. Voigt from University Children's Hospital Halle, M. Heiduk from University Hospital Magdeburg, U. Baumann and A. Zingel from Medical University Hannover, as well as N. Förster, A. Brunert, and S. Eckhardt from Children's Hospital Prinzessin Margaret Darmstadt).

Author's contributions: JW was project manager, responsible for data management, involved study design, and wrote the manuscript. DP was involved in study design, responsible for data analysis, and wrote the manuscript. TR, MKHA, HHU, MWL, PL, AK, NH, JdL, ACH, TK, GF, FS, and AR were pediatric gastroenterologists responsible as clinical investigators for recruitment of patients, gathered the clinical data and were involved in writing the manuscript. DH designed the study, was responsible for data analysis and wrote the manuscript. TM developed the idea for the study, was the principal investigator of the study, designed the study and wrote the manuscript.

Johannes Wolf, David Petroff, Dirk Hasenclever, and Thomas Mothes contributed equally to this work.

#### Conflicts of interest

These authors disclose the following: Thomas Mothes and Holm H. Uhlig were 2 of the inventors of the patent "Peptides and Their Use in a Procedure for Diagnostics of Coeliac Disease and Dermatitis Herpetiformis," (German patent DE10005932) with inventors' premiums paid by Leipzig University until 2014. Johannes Wolf received a grant from EUROIMMUN (Lübeck, Germany) for a celiac screening project in LIFE Child of the Research Centre of Civilization Diseases (Leipzig, Germany), outside the submitted work. Thomas Richter reports a grant from EUROIMMUN (Lübeck, Germany) for celiac screening in the Children's Hospital St Georg Leipzig, Germany, outside the submitted work. The remaining authors disclose no conflicts.

#### Funding

This study was funded by the European Regional Development Fund and an unrestricted grant from EUROIMMUN Medizinische Labordiagnostika AG (Dassow, Germany). EUROIMMUN Medizinische Labordiagnostika performed the blinded assays of antibodies and provided funding for the trial. EUROIMMUN played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of the manuscript.

## Details Regarding the Antibody Tests

The assays use a standard curve for IgA-TTG (anti-tissue transglutaminase enzyme-linked immunosorbent assay [ELISA] [IgA]; EA 1910 A) at 2, 20, and 200 U/mL and for IgG-DGL (anti-gliadin [GAF-3X]-ELISA [IgG], EV 3011 G) at 2, 25, 200 U/mL, as well as 2 controls (positive and negative control). For IgA-TTG, the coefficients of variation were 2.0% (intra-assay) and 5.2% (inter-assay). For IgG-DGL they were 5.3% (intra-assay) and 6.4% (inter-assay). For IgA-EMA testing on esophagus (FA 1911-1010 G), positive and negative controls are used. Each ELISA or EMA test batch passed a quality control showing the valid ranges for calibrators and controls. Both the ELISA and the EMA tests were produced according to the regulations for in vitro diagnostic medical device. Furthermore, tests are also validated regularly by ring trials.

## Statistical Details

### Diagnostic Tests With Indeterminate Diagnoses and Test Results

Estimates for diagnostic characteristics have to be adapted to our situation in which some patients have no diagnosis and our test procedures permit a “gray zone” that is neither test-positive nor test-negative. We use the notation  $D^+$ ,  $D^-$ , and  $D^0$  for positive, negative, and no final diagnosis, respectively, and  $T^+$ ,  $T^-$ , and  $T^0$  for the results of a test procedure. The symbol  $P(X|Y)$  denotes the conditional probability of  $X$  given  $Y$ . The notation for the number of patients in each category can be read off the contingency table.

	Test procedure negative ( $T^-$ )	Test procedure indeterminate ( $T^0$ )	Test procedure positive ( $T^+$ )	
Not-celiac disease ( $D^-$ )	$n_{11}$	$n_{12}$	$n_{13}$	$n_{D-}$
No final diagnosis ( $D^0$ )	$n_{21}$	$n_{22}$	$n_{23}$	$n_{D0}$
Celiac disease ( $D^+$ )	$n_{31}$	$n_{32}$	$n_{33}$	$n_{D+}$
	$n_{T-}$	$n_{T0}$	$n_{T+}$	$n$

Definition of diagnostic characteristics:	estimated by:
Sensitivity = $Se = P(T^+   D^+)$	$\widehat{Se} = n_{33} / n_{D+}$
Extended sensitivity = $eSe = P(T^+ \text{ or } T^0   D^+ \text{ or } D^0)$	$\widehat{eSe} = (n_{22} + n_{23} + n_{32} + n_{33}) / (n_{D+} + n_{D0})$
Specificity = $Sp = P(T^-   D^-)$	$\widehat{Sp} = n_{11} / n_{D-}$
Extended specificity = $eSp = P(T^- \text{ or } T^0   D^- \text{ or } D^0)$	$\widehat{eSp} = (n_{11} + n_{12} + n_{22} + n_{23}) / (n_{D-} + n_{D0})$

Prevalence =  $Pr = P(D^+)$   
 Extended prevalence =  $ePr = P(D^+ \text{ or } D^0)$

$$\widehat{Pr} = n_{D+} / n$$

$$\widehat{ePr} = (n_{D+} + n_{D0}) / n$$

If one assumes that these diagnostic properties are independent of prevalence, then Bayes' formula enables extrapolation from the observed PPVs and NPVs to situations with prevalence differing from the observed one.

$$PPV = P(D^+ | T^+) \quad \widehat{PPV} = \widehat{Se} \cdot \widehat{Pr} / [\widehat{Se} \cdot \widehat{Pr} + (1 - \widehat{eSp}) \cdot (1 - \widehat{Pr})]$$

$$NPV = P(D^- | T^-) \quad \widehat{NPV} = \widehat{Sp} \cdot (1 - \widehat{ePr}) / [\widehat{Sp} \cdot (1 - \widehat{ePr}) + (1 - \widehat{eSe}) \cdot \widehat{ePr}]$$

## Modifications With Changing Prevalence Due to Selection

Let us consider a target population for which we wish to draw inferences, for example, children and adolescents who visit the pediatric gastroenterologist and where a celiac disease (CD) antibody test is performed. To be concrete, let us assume this population has a 10% prevalence of CD. If one were to create a sample of this population with a desired prevalence by randomly choosing CD patients and non-CD patients in the predetermined proportion, then  $Se$ ,  $Sp$  of the sample would equal that of the entire population. However, the selection processes by which one arrives at a different prevalence in real life often differ substantially from random sampling. As a result, the condition “ $D^-$ ” in specificity, for example, can have quite a different meaning in the sample from the target population.

In our trial, many patients were selected for a biopsy because IgA-TTG concentrations in tests performed before the trial were in the vicinity of the manufacturer's cutoff. Thus, our sub-population with  $D^-$  has much higher IgA-TTG concentrations than from the average non-CD patient from the target population and, accordingly, specificity is lower. For example, specificity with the manufacturer's cut is only 89% in the selected AbCD population, but >97% in blood donors. This qualitative behavior holds for extended specificity as well.

We wish to accommodate this selection effect in our extrapolated estimates for PPV at lower prevalence. For NPV at low prevalence, failure to take this into account will be conservative, but so close to 100% that corrections are unnecessary. For PPV, corrections in estimates of  $Se$  will have minimal influence on the result and we need only take corrections for extended specificity ( $eSp$ ) into account. We do so by determining a value for  $eSp$  at low prevalence and then interpolating between this and the observed  $eSp$  value. First we discuss two approaches for obtaining estimates at low prevalence and then the method of interpolation.

## Estimates for Extended Specificity

### Model-Based Approach

With a small number of well-justified assumptions, we can use our data to model extended specificity at low prevalence.

We suppose that the target population of interest has a prevalence of about 10%, but note that our results do not depend strongly on this assumption.

We consider all not CD patients in the target population. Some are selected for the study population.

$P(\text{selected} | D^- \text{ or } D^0)$  can be calculated from the enrichment for CDs. Let PrevT denote the prevalence in the target population and PrevS in AbCD.

$$P(\text{selected} | D^- \text{ or } D^0) = (1 - \text{PrevS}) / \text{PrevS} * \text{PrevT} / (1 - \text{PrevT})$$

With PrevS = 59% and PrevT = 10% (5%, 15%) we get:  $P(\text{selected} | D^- \text{ or } D^0) = 7.7\%$  (3.7%–12.3%)

This is a strong selection leading to a major enrichment for CD cases in the study population.

We base the model considerations on the plausible assumption

- that an IgA-TTG pretest T1 is performed in the target population and
- all patients with values  $>1$  ULN are selected for the “trial” population, whereas those with lower values may or may not be selected.

By simple rules of probability:

Formula 1: Extended specificity in target population

$$= eSP(\text{selected}) \times P(\text{selected}) + eSP(\text{unselected}) \times (1 - P(\text{selected})) =$$

$$= 364/369 \times 0.077 + eSP(\text{unselected}) \times 0.923$$

Let  $T2^+$  denote a positive confirmatory IgA-aTTG test  $>10$  ULN, as above.

Formula 2:  $eSP(\text{unselected}) = 1 - P(T2^+ | (D^- \text{ or } D^0) \text{ and unselected})$ .

We use data from AbCD to derive an upper bound on  $P(T2^+ | (D^- \text{ or } D^0) \text{ and unselected})$ .

We use 2 pieces of information:

1. Our data indicate that 2 IgA-TTG tests performed on a single patient in short succession typically have a standard deviation of about a factor of 2 between them (cf. [Supplementary Figure 4](#)).

We model this intra-individual variability as follows: Each patient has an individual (latent) antibody expected value  $a$ . Measurements  $X$  on the  $\log_{10}$  scale vary normally around this mean with a standard deviation of  $SD_{rep}=0.3$  (see [Supplementary Figure 3](#)):

$$X \sim N(a, SD_{rep}^2).$$

Given the individual expected value, IgA-TTG measurements are independent.

This model implies that the predictive distribution of a second measurement given the (noninformative) posterior for the individual antibody expected value “ $a$ ” learned from a first measurement  $X1 = x1$  is:

$$(X2 | X1 = x1) \sim \text{Normal}(x1, 2*SD_{rep}^2).$$

Let  $\text{pnorm}(x, \text{mean}, \text{var})$  denote the cumulative normal distribution function:

$$\text{pr}(T2^+ | X1=x1) = \text{pr}(X2 > 1 | X1=x1) = 1 - \text{pnorm}(1, x1, 2*SD_{rep}^2),$$

remembering that the value  $X2 = 1$  on the  $\log_{10}$  scale corresponds to  $X2 = 10$  ULN.

2. The second piece of information we take from the data is the distribution of IgA-TTG measurements.

[Supplementary Figure 5](#) shows the distribution of IgA-aTTG measurements on the  $\log_{10}$  scale from the study population.

The distribution of IgA-aTTG measurements in the unselected part of the target population should be similar or shifted further to the left, because of selection for higher IgA-TTG measurements.

The red vertical line illustrates our assumption that patients with IgA-TTG  $>1$  ULN are all selected.

Thus it is conservative to take the observed data (figure 4) truncated to IgA-TTG  $\leq 1$  ULN as an approximation of the unknown pre-study IgA-TTG distribution in the unselected part of the target population.

$P(T2^+ | (D^- \text{ or } D^0) \leq \text{mean}(\text{pnorm}(1 | \text{obs}, 2*SD_{rep}^2))$  over all  $\text{obs} \leq 1$  ULN from not CD patients.

$$= 0.000185$$

Substituting in Formula 2 and 1 we get:

Extended specificity in target population =

$$= eSP(\text{selected}) * P(\text{selected}) + eSP(\text{unselected}) * (1 - P(\text{selected}))$$

$$\geq 364/369 \times 0.077 + (1 - 0.000185) \times 0.923$$

$$> \mathbf{0.9987}$$

For prevalence of  $<10\%$  the lower bound for eSP is even closer to one.

### Literature-Based Approach

In a recent retrospective study, Gidrewicz et al<sup>1</sup> present IgA-TTG data from 17,505 consecutive patients. Given the sensitivity and specificity of the test used (0.957 and 0.98, respectively), one can estimate that the prevalence of CD in their cohort is about 6%. Here too, the exact estimate of prevalence does not affect conclusions much. There were 336 patients with antibody concentrations  $>10$  ULN, 270 of which were followed up with biopsies. Given that 11 of these were diagnosed as not having CD, we can estimate that  $Sp$  (for  $10 \times$  ULN) in the entire cohort is about  $1 - 14/16390 = \mathbf{0.9991}$  in good agreement with the above estimate. As mentioned, this will hold roughly for eSP as well.



## Method of Interpolation

In the expression for PPV, we replace the estimate  $\widehat{eSp}$  by  $\widehat{eSp} = \widehat{eSp}_0 + (\widehat{eSp}_{\text{obs}} - \widehat{eSp}_0) \cdot \widehat{Pr} / \widehat{Pr}_{\text{obs}}$ , where the subscript “obs” refers to values observed in the study and “0” to the values at low prevalence. This expression concurs with the standard estimates for PPV at the observed prevalence and at low prevalence. Choosing a prevalence of zero for the low prevalence value is conservative, meaning we underestimate PPV. This method was used for PPV in Figure 3 in the Results section.

A linear interpolation of this sort is also conservative compared to a linear interpolation on the logit scale.

## Sensitivity Analyses

The effects of varying eSP are shown in Supplementary Figure 6. The optimistic scenario corresponds to having only 11 (eSP = 0.9993) instead of 14 patients in the Gidrewicz estimates who do not have CD, although antibody concentrations are >10 ULN. The pessimistic scenario corresponds to 20 such patients (eSP = 0.9988) and is the lower bound from the model-based approach. The very pessimistic scenario corresponds to 50 such patients (eSP = 0.9968). Even in the most pessimistic scenario, PPV is still reliable down to a prevalence of about 10%. Note that all our methods to determine PPV at low prevalence are conservative in multiple respects.

## Determination of lower confidence bounds

The LCB for PPV and NPV were determined from the logit transformed values using the delta method (eg, Mercaldo et al<sup>2</sup>). Here modifications are needed because of the 3 valued logic and because of the interpolations. We write the equations for the estimates, leaving off the hats.

$$\text{LCB(NPV)} = \exp(\text{logit(NPV)} - z_{1-\alpha} [\text{Var}(\text{logit(NPV)})]^{1/2}) / \{1 + \exp(\text{logit(NPV)} - z_{1-\alpha} [\text{Var}(\text{logit(NPV)})]^{1/2})\},$$

$$\text{LCB(PPV)} = \exp(\text{logit(PPV)} - z_{1-\alpha} [\text{Var}(\text{logit(PPV)})]^{1/2}) / \{1 + \exp(\text{logit(PPV)} - z_{1-\alpha} [\text{Var}(\text{logit(PPV)})]^{1/2})\},$$

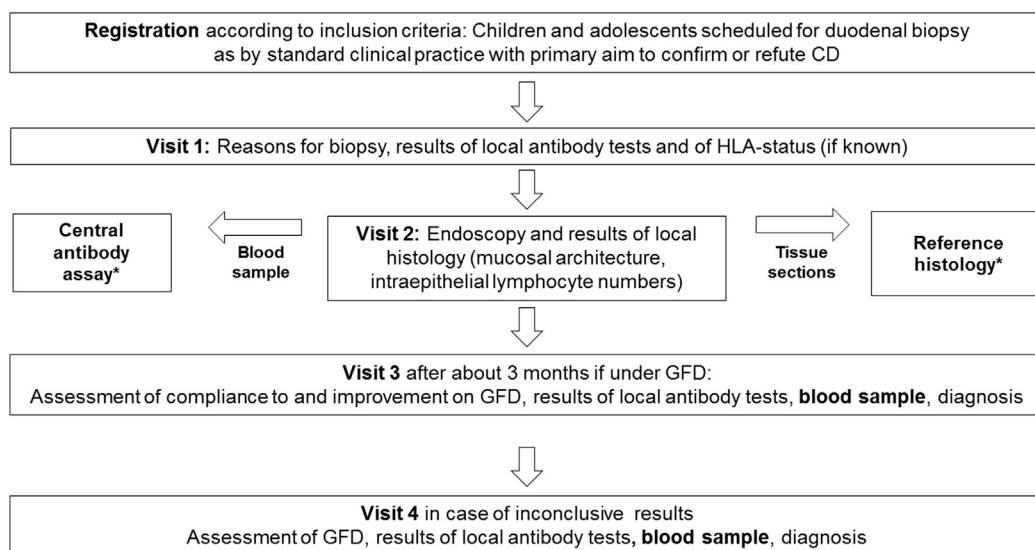
where  $\text{logit}(x) = \ln(x/(1-x))$  and  $z_{1-\alpha}$  denotes the  $(1-\alpha)$  quantile of the inverse normal distribution. The variance terms are

$$\text{Var}(\text{logit(NPV)}) = (1 - \text{Sp}) / ((n - n_{D+}) \cdot \text{Sp}) + \text{eSe} / (n_{D+} \cdot (1 - \text{eSe}))$$

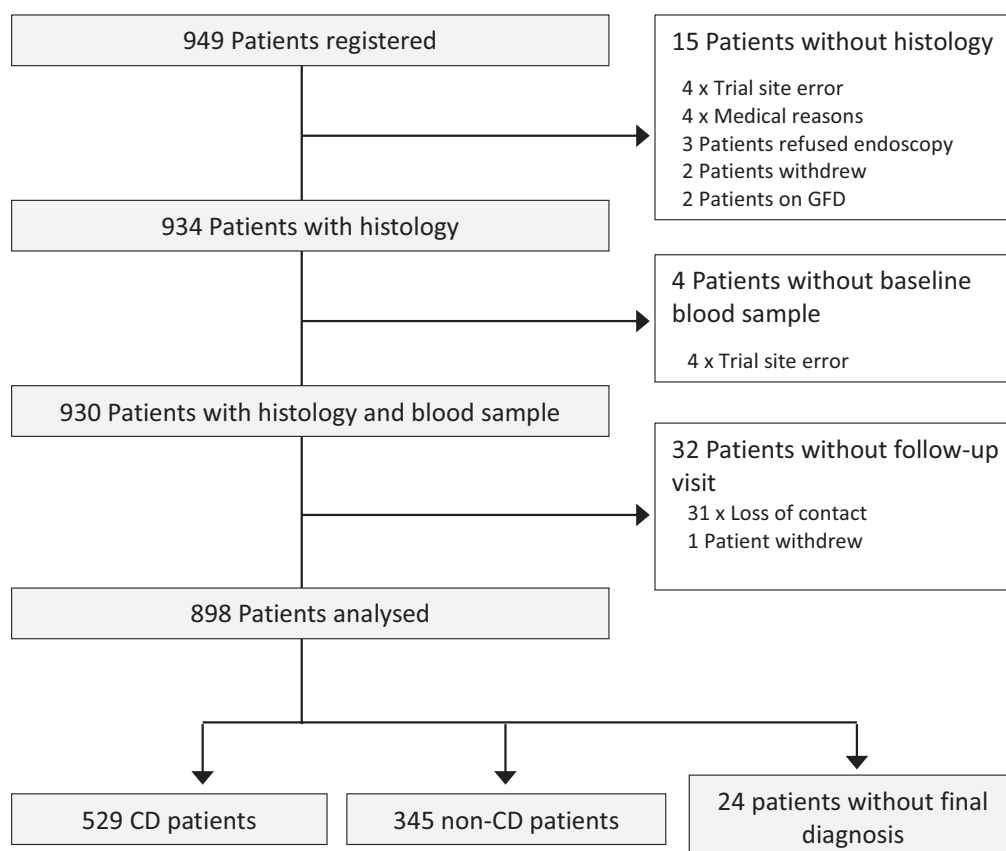
$$\text{Var}(\text{logit(PPV)}) = (1 - \text{Se}) / (n_{D+} \cdot \text{Se}) + \text{eSp}_{\text{obs}} \cdot (1 - \text{eSp}_{\text{obs}}) / (1 - \text{eSp})^2 \cdot (\text{Pr} / \text{Pr}_{\text{obs}})^2 / (n - n_{D+}) + \text{eSp}_0 \cdot (1 - \text{eSp}_0) / (1 - \text{eSp})^2 \cdot (1 - \text{Pr} / \text{Pr}_{\text{obs}})^2 / n_0$$

## References

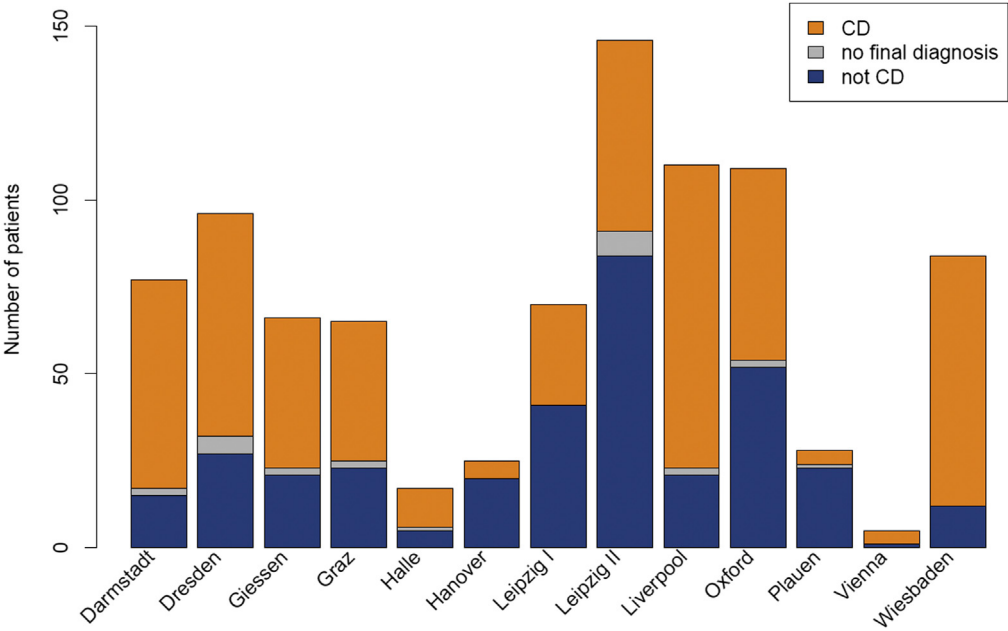
1. Gidrewicz D, Potter K, Trevenen CL, et al. Evaluation of the ESPGHAN Celiac Guidelines in a North American Pediatric population. *Am J Gastroenterol* 2015;110:760–767.
2. Mercaldo ND, Lau KF, Zhou XH. Confidence intervals for predictive values with an emphasis to case-control studies. *Stat Med* 2007;26:2170–2183.



**Supplementary Figure 1.** Simplified study procedure. \*Feedback was given after GFD decision (no later than 3 months after inclusion), but only if blood, histology slides, and all data of visit 1 and visit 2 were available.

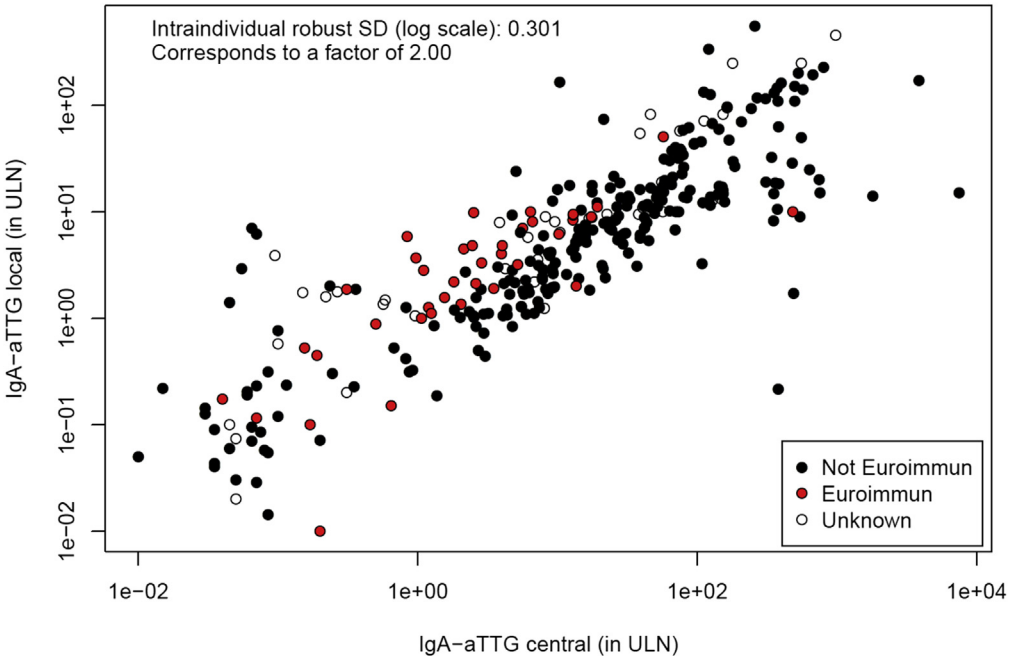


**Supplementary Figure 2.** Trial recruitment.

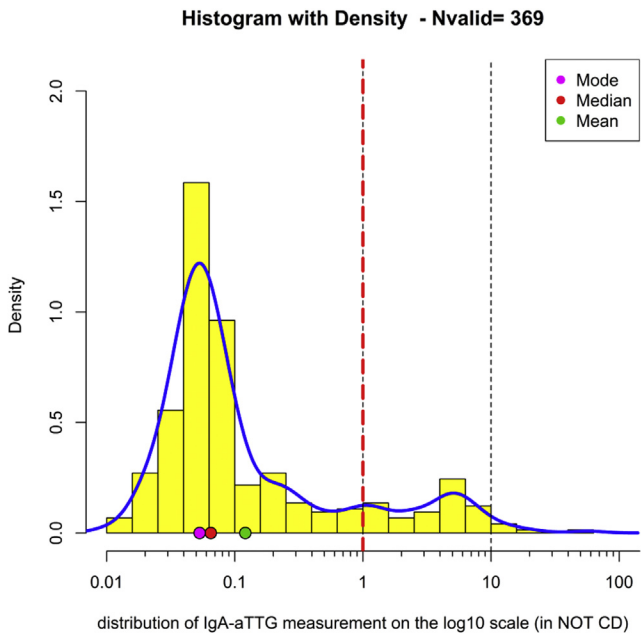


**Supplementary Figure 3.** Distribution of patient accrual and diagnoses by trial site. Leipzig I, University Children's Hospital Leipzig, Germany; Leipzig II, Children's Hospital of the Clinical Centre "Sankt Georg" Leipzig, Germany.

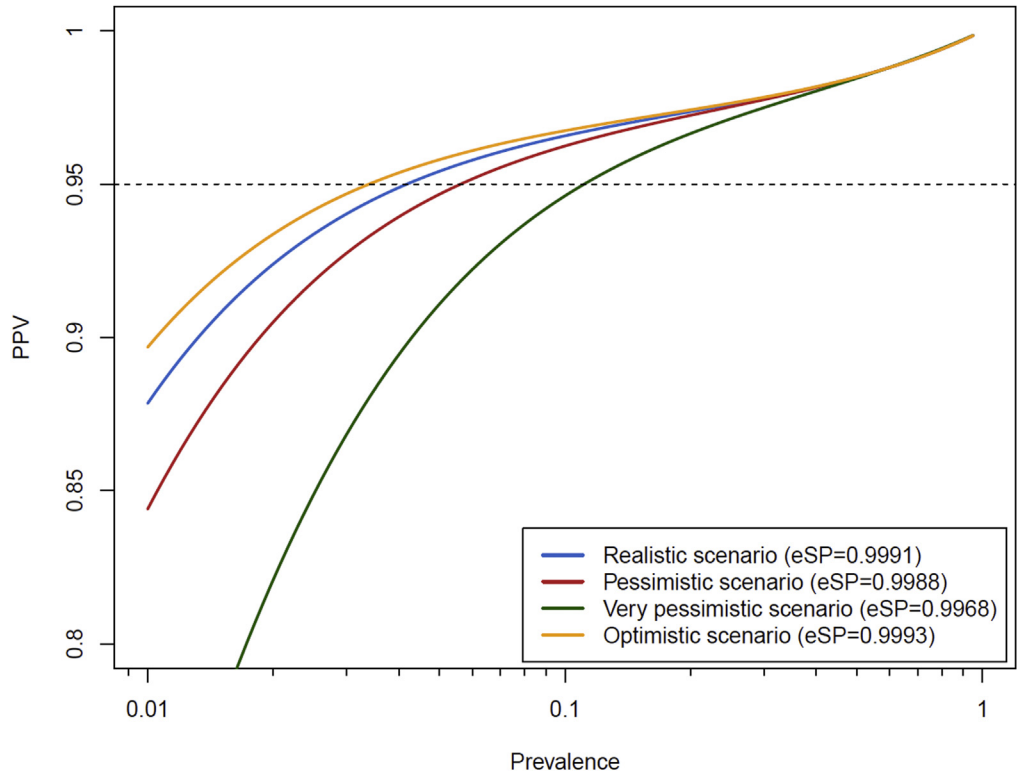
**Comparison of all tests (duplicates removed)**



**Supplementary Figure 4.** Intra-individual variance. The last available local IgA-TTG concentration is plotted against the value measured in the central laboratory in units of ULN. More than half the tests were performed within 1 week of each other. The intra-individual SD is provided using a robust estimate based on median absolute deviations.



**Supplementary Figure 5.** Antibody distribution for non-CD patients.



**Supplementary Figure 6.** Sensitivity Analyses.



**Supplementary Table 1.** Symptoms and Signs Suggestive for Celiac Disease in Patients of the Trial at Baseline

Variable	All (n = 898)	CD (n = 529)	Not CD (n = 345)	NFD (n = 24)
Prior positive antibody test	603 (67)	496 (94)	87 (25)	20 (83)
Abdominal pain	547 (61)	295 (56)	242 (70)	10 (42)
Diarrhea	251 (28)	148 (28)	91 (26)	12 (50)
Bloating	170 (19)	113 (21)	48 (14)	9 (38)
Constipation	130 (14)	76 (14)	51 (15)	3 (12)
Cramping	129 (14)	65 (12)	59 (17)	5 (21)
Failure to thrive	120 (13)	71 (13)	42 (12)	7 (29)
Nausea <sup>a</sup>	113 (13)	37 (7)	72 (21)	4 (17)
Fatigue	111 (12)	73 (14)	37 (11)	1 (4)
Vomiting <sup>a</sup>	106 (12)	36 (7)	66 (19)	4 (17)
Weight loss	106 (12)	64 (12)	37 (11)	5 (21)
Flatulence	99 (11)	61 (12)	37 (11)	1 (4)
Iron deficiency anemia <sup>a</sup>	76 (8)	56 (11)	20 (6)	0 (0)
Short stature	75 (8)	52 (10)	23 (7)	0 (0)
Anorexia	37 (4)	18 (3)	17 (5)	2 (8)
Apthous stomatis	22 (2.4)	13 (2.5)	9 (2.6)	0 (0)
Dermatitis herpetiformis-like rash	12 (1.3)	10 (1.9)	2 (0.6)	0 (0)
Abnormal liver chemistry	10 (1.1)	8 (1.5)	2 (0.6)	0 (0)
Delayed puberty	4 (0.4)	3 (0.6)	1 (0.3)	0 (0)
Amenorrhea	3 (0.3)	2 (0.4)	1 (0.3)	0 (0)
Osteopenia/osteoporosis	1 (0.1)	0 (0)	1 (0.3)	0 (0)

NOTE. Data are n (%).

NFD, no final diagnosis.

<sup>a</sup>Significant differences between CD and control patients ( $P < .05$ ; Fisher's exact probability test).**Supplementary Table 2.** List of Antibody Tests Kits Used Locally at the Recruitment Centers

Center	IgA-TTG	IgG-DGL
Leipzig I (Germany)	Inova Quanta Lite	Inova Quanta Lite
Leipzig II (Germany)	Phadia EliA Celikey	Phadia EliA Celikey
Dresden (Germany)	Generic Assay/Inova Quanta Flash	—
Giessen (Germany)	EUROIMMUN	EUROIMMUN
Wiesbaden (Germany)	Phadia EliA Celikey	Phadia EliA Celikey
Vienna (Austria)	Orgentec	—
Graz (Austria)	Eurospital	—
Oxford (United Kingdom)	Inova Quanta Flash	—
Liverpool (United Kingdom)	Phadia EliA Celikey	—
Halle (Germany)	Phadia EliA Celikey	Phadia EliA Celikey
Plauen (Germany)	EUROIMMUN	—
Darmstadt (Germany)	Phadia EliA Celikey	Phadia EliA Celikey
Hanover (Germany)	Unknown	—

**Supplementary Table 3.** Comparison of Reference and Local Diagnostic Procedures

Variable	TTG-IgA procedure <sup>a</sup>			TTG-DGL procedure <sup>b</sup>		
	Reference (n = 882)	Local antibody measurement		Reference (n = 898)	Local antibody measurement	
		Not >7 d before trial blood sample (n = 326)	All measurements (n = 682)		Not >7 d before trial blood sample (n = 154)	All measurements (n = 321)
PPV, estimate (95% LCB)	0.988 (0.975)	0.984 (0.952)	0.989 (0.972)	0.988 (0.975)	0.963 (0.891)	0.983 (0.948)
NPV, estimate (95% LCB)	0.934 (0.908)	0.899 (0.844)	0.908 (0.870)	0.958 (0.934)	0.948 (0.874)	0.948 (0.895)
Proportion of CD patients >10× ULN	0.772	0.618	0.624	0.766	0.667	0.652
Proportion of patients to be biopsied <sup>c</sup>	0.173	0.273	0.295	0.203	0.273	0.333

NOTE. For list of antibody test kits applied in the recruitment centers, see [Supplementary Table 2](#). Local antibody results were only used for calculations if quantitative data were available (but not if prior test results were expressed as “positive,” “negative,” or “unclear”) and the test manufacturer together with the cutoff was known. In 2 of the centers, the EUROIMMUN test kits were also applied for local antibody measurements. For calculation of PPV and NPV, patients without final diagnosis were not excluded and thus counted as false positive and false negative, respectively (see also [Table 3](#)).

<sup>a</sup>Children with sIgAD were excluded before analysis.

<sup>b</sup>Children with sIgAD were not excluded before analysis.

<sup>c</sup>Patients to be biopsied are children with antibody values in the gray zone (1× ULN to 10× ULN).

**Supplementary Table 4.** Discrepancies Between Final Histology and Final Diagnosis

No.	Patient code	Age (years)/sex	Symptoms and associated disorders <sup>a</sup>	Initial				HLA	Histology local	Histology final	Follow-up		Remarks and differential diagnoses
				IgA-TTG local	IgA-TTG central	IgG-DGL central	IgA-EMA central				IgA-TTG local	IgA-TTG central	
1	05997 Non-CD	16/M	1,2,8 slgAD	0.1	0.0	0.1	Negative	DQ2	Normal	Marsh 3A	NM	NM	Symptoms disappear despite gluten exposure
2	43342 Non-CD	13/M	2,8	0.0	0.0	0.0	Negative	NM	Normal	Marsh 3A	NM	NM	<i>Helicobacter pylori</i> infection Symptoms disappear despite gluten exposure
3	98257 Non-CD	3/M	2,8,9,10	0.0	0.0	3.1	Negative	NM	Marsh 1	Marsh 3A	NM	NM	Eosinophilic esophagitis Suspicion of cow's milk allergy
4	57176 Non-CD	1/F	3,12,13	NM	0.0	0.4	Negative	DQ8	Marsh 3A	Marsh 3A	0.1	0.1	No improvement under GFD
5	66528 Non-CD	1/F	12,13	0.0	0.0	0.4	Negative	Negative	Marsh 3A	Marsh 3A	NM	NM	Wheat allergy Improvement under GFD
6	80741 Non-CD	14/F	3,8,9	0.0	0.0	0.0	Negative	NM	Marsh1	Marsh 3A	NM	NM	Undefined gastritis
7	93926 Non-CD	1/M	1,3	0.0	0.0	0.0	Negative	DQ8	Marsh 3A	Marsh 2	0.0	0.0	Iron deficiency Improvement of symptoms under normal diet
8	69612 Non-CD	15/M	1,2,3,9	NM	0.1	0.2	Negative	Negative	Normal	Marsh 3B	NM	NM	Crohn's disease
9	16571/ Non-CD	16/F	2,5,7,9,11	NM	0.1	0.0	Negative	Negative	Marsh 1	Marsh 3A	NM	NM	Psychosomatic disorder
10	82817 Non-CD	17/F	1,2,7	NM	0.1	0.1	Negative	NM	Normal	Marsh 3A	0.1	0.0	Fructose and sorbitol intolerance, proctitis Diarrhea stopped, constipation persisted
11	61902 Non-CD	1/F	2,12,13	0.1	0.2	0.1	Negative	NM	Marsh 3A	Marsh 3B	NM	NM	No improvement under GFD
12 <sup>b</sup>	27680 Non-CD	7/M	2,4	Negative	0.6	0.6	Negative	DQ2	Marsh 3B	Marsh 3B	1.2	1.2	Growth hormone deficiency Fluctuating antibodies for 3 years
13	03784 Non-CD	14/F	2,5	Positive	2.5	0.4	Negative	DQ2	Normal	Marsh 3A	1.3	0.4	Gastric ulcer
14	50139 CD	5/F	1,2,3,4, 5,6,9	3.9	0.1	0.3	Negative	DQ2	Patchy Marsh 3A	Normal	0.0	0.7	Improvement under GFD
15 <sup>b</sup>	00521 CD	6/M	1,2,3,6	2.0	0.2	0.0	Negative	DQ2	Normal	Normal	NM	NM	Partial GFD between visit 1 and 2 not fully excluded. Clinical remission under GFD
16	06940 CD	6/F	1,2,3,4, 5,11	2.8	4.7	0.6	1:32	DQ2, DQ8	PVA	PVA	2.5	0.8	Clinical and serologic remission under GFD

Supplementary Table 4. Continued

No.	Patient code	Age (years)/sex	Symptoms and associated disorders <sup>a</sup>	Initial				HLA	Histology local	Histology final	Follow-up		Remarks and differential diagnoses
				IgA-TTG local	IgA-TTG central	IgG-DGL central	IgA-EMA central				IgA-TTG local	IgA-TTG central	
17	94157 CD	4/F	2,14	3.1	7.5	0.8	1:100	NM	Normal	Normal	8.2 <sup>c</sup>	0.2 <sup>c</sup>	Clinical and serologic remission under GFD (recommended at visit 3)
18	67550 CD	15/F	None	6.1	29.7	1.0	1:320	DQ2	Patchy Marsh 3A	Normal	5.9	0.5	Clinical and serologic remission under GFD

NOTE. Antibody assays and histology performed in local trial centers and centrally. Concentrations of IgA-TTG and IgG-DGL expressed as multiples of cutoffs ( $\times$ ULN). The result of the final central histology inspection is shown. In 13 cases (nos. 1–13) CD was excluded, although final histology was CD compatible. In all but 1 (no. 12), CD was excluded on the basis of HLA incompatibility, normal or Marsh 1 local histology, differential diagnoses, lack of improvement under normal diet or improvement under GFD. In 5 cases (nos. 14–18) CD was diagnosed although final histology excluded CD. In all, CD was diagnosed on the basis of improvement under GFD. In 3 cases (nos. 13, 16, and 17), antibody positivity is in the gray zone between  $1\times$ ULN and  $10\times$ ULN, therefore, these patients are excluded from the diagnostic algorithms. All except case 1 are IgA competent.

F, female; M, male; NA, not applicable; NM, not measured; PVA, partial villous atrophy; T1DM, diabetes mellitus type 1.

<sup>a</sup>Symptoms: 1 = diarrhea; 2 = abdominal pain; 3 = abdominal distension; 4 = flatulence; 5 = cramping; 6 = fatigue; 7 = short stature; 8 = vomiting; 9 = constipation; 10 = weight loss; 11 = nausea; 12 = anorexia; 13 = failure to thrive; 14 = gastroesophageal reflux.

<sup>b</sup>In 2 patients (12 and 15), an influence of the antibody data on diagnosis or exclusion of CD may have been relevant.

<sup>c</sup>Measured under gluten intake.



**Supplementary Table 5.** Comparison of First Reference Histologic Evaluation With Results of the Two Antibody Procedures

Variable	Classification according to TTG-IgA procedure <sup>a,b</sup>			Classification according to TTG-DGL procedure <sup>c</sup>		
	Non-CD (n = 320)	Indeterminate (n = 153)	CD (n = 409)	Non-CD (n = 306)	Indeterminate (n = 182)	CD (n = 410)
No assessment possible <sup>d</sup>	0	2	0	0	2	0
IEL not elevated						
Histology normal	258	31	5	252	45	5
Crypt hyperplasia	8	7	0	9	7	0
Partial villous atrophy	4	2	1	4	2	1
Subtotal villous atrophy	2	0	0	1	1	0
Total villous atrophy	0	0	0	0	0	0
IEL elevated						
Marsh 1	20	3	0	17	6	0
Marsh 2	7	4	7	6	6	7
Marsh 3A	15	47	64	12	52	64
Marsh 3B	6	48	188	5	52	188
Marsh 3C	0	9	144	0	9	145

<sup>a</sup>The TTG-IgA procedure does not include patients with sIgAD.

<sup>b</sup>Classification according to TTG-IgA procedure: CD if IgA-TTG  $\geq 10 \times$  ULN, non-CD if CD if IgA-TTG  $< 1 \times$  ULN, otherwise indeterminate.

<sup>c</sup>Classification according to TTG-DGL procedure: CD if IgA-TTG  $\geq 10 \times$  ULN or IgG-DGL  $\geq 10 \times$  ULN; non-CD if IgA-TTG  $< 1 \times$  ULN and IgG-DGL  $< 1 \times$  ULN, otherwise indeterminate.

<sup>d</sup>Insufficient quality of inspected tissue, no information on mucosal architecture or IEL number.

**Supplementary Table 6.** Patients With False-Positive Results for the Transglutaminase Against IgA and Transglutaminase Against Deamidated Gliadin Procedure

No.	Patient code	Age, y/sex	Symptoms	Associated disorders	Initial				HLA	Histology local	Histology final	Follow-up	
					IgA-TTG local	IgA-TTG central	IgG-DGL central	IgA-EMA central				IgA-TTG central	IgG-DGL central
1	27591 Non-CD	13/F	None	T1DM	10.3	14.7	0.3	1:320	DQ8	Normal	Marsh 1	9.6	0.3
2	54242 Non-CD	11/M	Abdominal pain (disappear despite gluten exposure)	JIRA	9.3	12.9	0.8	1:100	DQ2	Normal	Normal	2.6	0.6
3	60633 NFD	8/M	Cramping, abdominal pain	None	9.9	56.6	0.4	1:320	DQ2	Normal	Normal	NM	NM
4	29996 NFD	2/F	Failure to thrive, weight loss, cramping, abdominal distension, abdominal pain (disappear despite gluten exposure)	None	2.9	21.2	2.1	1:100	DQ2	Marsh 2	Marsh 2	34.2	0.6
5	48427 NFD	14/M	Abdominal pain, diarrhea, lactose intolerance (disappear despite gluten exposure)	None	2.6	11.6	4.6	1:100	DQ2	Normal	Marsh 3A	8.4	4.9

NOTE. Patients without final diagnosis included in this table because they are considered as false positives for calculation of PPV. Antibody assays and histology performed in local trial centers and centrally. Concentration of IgA-aTTG and of IgG-aDGL expressed as multiples of cutoff ( $\times$ ULN). The result of the final central histology inspection is shown.

F, female; JIRA, juvenile idiopathic rheumatic arthritis; M, male; NFD, no final diagnosis; NM, NM not measured; T1DM, type 1 diabetes mellitus.

**Supplementary Table 7.** IgA-Competent Patients False Negative for IgA Against Transglutaminase

Patient code	Age, y/sex	Symptoms <sup>a</sup> and associated disorders	Duration of complaints	Initial					HLA	Histology local	Histology final	Follow-up		Improvement under GFD	Remarks
				IgA-TTG local	IgA-TTG central	IgG-DGL central	IgA-EMA central	IgA-TTG central				IgG-DGL central			
01119 CD	5/M	1,2,3,4	7 mo	NM	0.0	0.5	Negative	DQ2	Marsh 3A	Marsh 3A	0.1	0.4	Yes	None	
05325 CD	2/F	3,9	Unknown	0.0	0.0	1.0	Negative	DQ2	Marsh 3A	Marsh 3B	0.0 <sup>b</sup>	1.7 <sup>b</sup>	Yes	Due to persistent complaints GFD recommended at visit 3	
58339 <sup>c</sup> CD	5/F	2,9	4 y	0.1	0.0	1.3	Negative	DQ2	Marsh 1	Marsh 3A	NM	NM	Yes	Due to persistent complaints GFD recommended at visit 3	
65855 CD	12/M	First-degree relation with T1DM	NA	7.0	0.1	0.1	Negative	NM	Marsh 3A	Marsh 3B	3.9	0.1	NA	Sample mix-up claimed to be excluded by investigator for the initial blood sample	
50139 <sup>c</sup> CD	5/F	1,2,3,4,5,6	5 y	3.9	0.1	0.3	Negative	DQ2	Marsh 3A	Normal	0.0	0.7	Yes	Local histology reported as patchy	
2728X CD	4/F	1,3,4,7,8, Down syndrome	3 y	0.0	0.1	0.1	Negative	DQ2	Marsh 3A	Marsh 2	NM	NM	Yes	None	
00895 CD	5/M	2,3, Down syndrome, hypothyroidism	Unknown	0.4	0.2	3.6	Negative	DQ2, DQ8	Marsh 3C	Marsh 3B	0.0	1.6	Yes	None	
00521 CD	6/M	1,2,3	1 y	2	0.2	0.0	Negative	DQ2	Normal	Normal	NM	NM	Yes	Partial GFD between visit 1 and 2 not fully excluded	
99557 CD	8/M	2,9	7 y	0.5	0.7	0.5	Negative	NM	Marsh 3A	Marsh 3B	0.3	0.3	Yes	IgA-TTG and IgG-DGL slightly positive in a prior test of the general practitioner	
71773 CD	8/F	5,9	7 y	0.4	0.8	1.3	negative	NM	Marsh 3A	Marsh 3B	0.4	0.8	Yes	None	
21419 CD	16/F	1,2,3,5,10, autoimmune thyroiditis, partial IgA deficiency <sup>d</sup>	4 mo	0.3	0.9	4.0	1:10	DQ2	Marsh 3A	Marsh 3A	0.4	1.0	Yes	IgG-EMA and IgG-TTG positive	

Supplementary Table 7. Continued

Patient code	Age, y/sex	Symptoms <sup>a</sup> and associated disorders	Duration of complaints	Initial					HLA	Histology local	Histology final	Follow-up		Improvement under GFD	Remarks
				IgA-TTG local	IgA-TTG central	IgG-DGL central	IgA- EMA central	IgA-TTG central				IgG-DGL central			
31886 CD	7/M	1,2,5,10,11, aphtous stomatis	14 mo	0.3	0.9	1.7	Negative	DQ2	Marsh 3A	Marsh 2	0.2	1.9	Yes	None	
10937 CD	10/M	2,10	3 mo	Positive	0.9	0.3	Negative	NM	Marsh 3C	Marsh 3A	0.1	0.7	Yes	None	
25135 CD	13/F	6,9, dermatitis herpetiformis-like rash	7 mo	1.0	0.96	0.3	Negative	NM	Marsh 3A	Marsh 2	NM	NM	Yes	None	
06063 CD	7/F	2	5 y	3.6	0.97	0.7	Negative	NM	Marsh 3B	Marsh 3B	0.5	0.7	Yes	None	
58952 NFD	1/M	1,3,8,12,13, partial IgA deficiency <sup>d</sup>	1.5 y	0.0	0.0	0.0	Negative	NM	Marsh 3A	Marsh 3A	0.0	0.0	Yes	Symptoms resolving partially under normal diet	
01660 <sup>c</sup> NFD	2/M	13	Unknown	1.4	0.0	0.3	Negative	NM	Normal	Marsh 2	NM	NM	NA	Initial local IgG-DGL = 1.9	
85490 <sup>c</sup> NFD	13/F	8,11	Unknown	0.0	0.0	1.7	Negative	Negative	Marsh 3A	Marsh 1	7.5	2.0	No	Initial local IgG-DGL=1.4	
87881 NFD	1/M	9,10,13	1.3 y	0.0	0.1	0.3	Negative	DQ2	Marsh 1	Marsh 1	5	0.0	Yes	Partial GFD between visit 1 and 2 not fully excluded	
03383 NFD	1/M	8,11	1.5 y	Negative	0.2	1.1	Negative	NM	Marsh 1	Normal	0.2	0.1	Yes	Follow-up IgA-EMA=1:100	
13442 NFD	6/M	1,2,3,4,6,10, first-degree relation with T1DM	4 mo	0.9	0.5	0.0	Negative	DQ8	Normal	Normal	0.6	0.0	NA	Local IgA-DGL increase from 4×ULN to 12.2×ULN	
														Symptoms not resolving under normal diet	

NOTE. Patients without final diagnosis included in this table because they are considered as false negatives for calculation of NPV. Antibody assays and histology performed in local trial centers and centrally. Concentration of IgA-TTG and IgG-DGL expressed as multiples of cutoff (×ULN). The result of the final central histology inspection is shown.

F, female; M, male; NA, not applicable; NFD, no final diagnosis; NM, not measured; T1DM, type 1 diabetes mellitus.

<sup>a</sup>Symptoms: 1 = diarrhea; 2 = abdominal pain; 3 = abdominal distension; 4 = flatulence; 5 = cramping; 6 = fatigue; 7 = short stature; 8 = vomiting; 9 = constipation; 10 = weight loss; 11 = nausea; 12 = anorexia; 13 = failure to thrive.

<sup>b</sup>Measured under gluten intake.

<sup>c</sup>Four patients with major discrepancies between local and final histology.

<sup>d</sup>Partial IgA-deficiency; IgA between 0.07 g/L and the lower age-dependent cutoff.



Supplementary Table 8. Patients Without Final Diagnosis

Patient code	Age, y/sex	Symptoms, <sup>a</sup> associated disorders and conditions	Initial				HLA	Histology local	Histology final	Follow-up		Remarks
			IgA-TTG local	IgA-TTG central	IgG-DGL central	IgA-EMA central				IgA-TTG central	IgG-DGL central	
58952	1/M	1,3,8,12,13, partial IgA deficiency	0.0	0.0	0.1	Negative	NM	Marsh 3A	Marsh 3A	0.0	0.0	Symptoms partially resolve under GFD
01660	2/M	13	1.4	0.1	0.3	Negative	NM	Normal	Marsh 3A	NM	NM	No GFD
87881	1/M	9,10,13	0.0	0.1	0.3	Negative	DQ2	Normal	Marsh 1	5.0 <sup>b</sup>	0.0 <sup>b</sup>	Partial GFD between visit 1 and 2 not fully excluded. Symptoms resolving under GFD. IgA-EMA at follow-up = 1:100
85490	13/F	8,11	0.0	0.1	1.7	Negative	negative	Marsh 3A	Marsh 1	7.5	2.0	Symptoms not resolving under GFD
03383	1/M	8,11	Negative	0.3	1.1	Negative	NM	Marsh 3A	Normal	0.2	0.1	Symptoms resolve under GFD
13442	6/M	1,2,3,4,6,10, first-degree relation with T1DM	0.9	0.5	0.0	Negative	DQ8	Normal	Normal	0.6 <sup>b</sup>	0.0 <sup>b</sup>	Local IgA-DGL increase from 4× ULN to 12.2× ULN. Symptoms not resolving under normal diet
10762	4/F	2,9,13	2.8	1.1	0.0	1:10	DQ2	Normal	CH	0.0	0.0	Symptoms not resolving under GFD
28994	6/M	1,2	0.9	1.3	0.9	1:10	DQ2	Marsh 1	Marsh 3A	NM	NM	Symptoms not resolving under normal diet
03767	13/F	1,2,3,5,11	0.8	2.6	0.7	1:10	DQ2	Marsh 3B	Marsh 3A	NM	NM	Fructose and lactose intolerance
21186	15/F	2,3,10	Positive	4.1	0.5	Negative	NM	Marsh 1	Normal	2.5	0.4	Symptoms resolving under GFD (GFD on patient's initiative)
29605	3/M	1,3,13	1.1	4.1	0.6	1:10	DQ2	Marsh 3B	Marsh 3B	NM	NM	Hypopituitarism, hypercortisolism. Symptoms not resolving under normal diet
54910	7/M	1,3, first-degree relation with CD	0.8	4.8	0.7	1:32	DQ2	Marsh 3A	Marsh 3B	3.1 <sup>b</sup>	0.7 <sup>b</sup>	Symptoms not resolving under normal diet
8676X	9/M	8,11	2.1	5.0	2.2	1:32	DQ2, DQ8	Normal	NAP, 25 IEL/100 enterocytes	7.5 <sup>b</sup>	0.7 <sup>b</sup>	Symptoms resolve under normal diet
96061	9/M	TDM1	6.9	5.6	0.3	1:32	DQ2	Normal	Normal	10.6 <sup>b</sup>	0.4 <sup>b</sup>	No GFD
98702	13/M	TDM1	1.8	5.7	1.5	1:100	DQ2	Normal	Marsh 3A	4.6 <sup>b</sup>	1.7 <sup>b</sup>	No GFD
61652	5/F	1,2,3,9, first-degree relation with CD	1.1	6.0	1.0	1:32	DQ2, DQ8	Normal	Normal	NM	NM	Symptoms not resolving under normal diet
92674	2/M	1, endogenous eczema	10.0	6.6	1.4	1:100	DQ8	NAP	Marsh 3A	NM	NM	Local IgA-TTG=6.0× ULN at follow-up. Diarrhea not resolving under normal diet
67043	10/M	1,5,10,13	2.4	7.7	0.3	1:100	DQ8	Normal	Normal	NM	NM	Symptoms resolve under normal diet
5109X	3/M	1,2 T1DM	2.3	7.8	1.5	1:320	DQ2, DQ8	Normal	Normal	3.8 <sup>b</sup>	0.9 <sup>b</sup>	Symptoms not resolving under normal diet
52221	4/F	5, stool abnormalities	2.0	8.3	1.7	1:100	DQ8	Marsh 1	Marsh 1	3.1 <sup>b</sup>	2 <sup>b</sup>	Symptoms not resolving under normal diet

Supplementary Table 8. Continued

Patient code	Age, y/sex	Symptoms, <sup>a</sup> associated disorders and conditions	Initial				HLA	Histology local	Histology final	Follow-up		Remarks
			IgA-TTG local	IgA-TTG central	IgG-DGL central	IgA-EMA central				IgA-TTG central	IgG-DGL central	
03232	1/F	1, first-degree relation with T1DM	2.6	9.2	3.0	1:100	DQ2	Normal	Normal	1.9 <sup>b</sup>	1.5 <sup>b</sup>	Symptoms not resolving under normal diet
48427	14/M	1,2, lactose intolerance	2.6	11.6	4.6	1:100	DQ2	Normal	Marsh 3A	8.4 <sup>b</sup>	4.9 <sup>b</sup>	Symptoms resolve under normal diet
29996	2/F	2,3,5,10,13	2.9	21.2	2.1	1:100	DQ2	Marsh 2	Marsh 2	34.2 <sup>b</sup>	0.5 <sup>b</sup>	Symptoms disappear despite gluten exposure
60633	8/F	2,5	9.9	56.6	0.4	1:320	DQ2	Normal	Normal	NM	NM	Symptoms not resolving under normal diet

NOTE. Antibody assays and histology performed in local trial centers and centrally. Concentration of IgA-TTG and of IgG-DGL expressed as multiples of cutoffs ( $\times$ ULN). The result of the final central (consensus) histology inspection is shown.

CH, crypt hyperplasia; NA, not applicable; NAP, no assessment of mucosal architecture possible; NM, not measured; T1DM, diabetes mellitus type 1.

<sup>a</sup>Symptoms: 1 = diarrhea; 2 = abdominal pain; 3 = abdominal distension; 4 = flatulence; 5 = cramping; 6 = fatigue; 7 = short stature; 8 = vomiting; 9 = constipation; 10 = weight loss; 11 = nausea; 12 = anorexia; 13 = failure to thrive.

<sup>b</sup>Measured under gluten-intake.

**Supplementary Table 9.** Reasons for Inclusion of Asymptomatic Patients Into the Study

Reasons for inclusion	All (n = 74)	CD patients (n = 59)	Non-CD patients (n = 14)	No final diagnosis (n = 1)
Associated disorders <sup>a</sup>	35 (47)	27 (46)	7 (50)	1 (100)
+ prior antibody positivity				
First-degree relatives with CD	22 (30)	21 (36)	1 (7)	0
+ prior antibody positivity				
First-degree relatives with CD	3 (4)	3 (5)	0	0
+ associated disorders <sup>b</sup>				
+ prior antibody positivity				
First-degree relation with diabetes mellitus type 1	2 (3)	2 (3)	0	0
+ prior antibody positivity				
Prior antibody positivity	3 (4)	3 (5)	0	0
+ HLA-compatibility				
Prior antibody positivity	5 (7)	3 (5)	2 (14)	0
First-degree relative with CD	2 (3)	0	2 (14)	0
or associated disorder <sup>c</sup>				
Other reasons <sup>d</sup>	2 (3)	0	2 (14)	0

NOTE. Values are n (%). Prior antibody positivity means positive results of IgA-TTG or IgA-EMA tests in the recruitment centers before inclusion into the trial.

<sup>a</sup>Thirty-three patients with type 1 diabetes mellitus and 2 with autoimmune thyroiditis.

<sup>b</sup>All with type 1 diabetes mellitus.

<sup>c</sup>One with IgA-deficiency and HLA-DQ 2.

<sup>d</sup>One patient with positive prior IgA-DGL test and 1 with sacroiliitis and elevated calprotectin.