Intestinal biopsy to diagnose celiac disease is not always required:

Importance of combined antibody tests

Bürgin-Wolff Annemarie\textsuperscript{1}, Buser Mauro\textsuperscript{2}, Hadziselimovic Faruk\textsuperscript{1*}

Institute for Celiac disease, Liestal, Switzerland\textsuperscript{1}, Statistik Dr. M. Buser, Riehen, Switzerland\textsuperscript{2}
ABSTRACT

Objective: To compare celiac disease (CD)–specific antibody tests to determine if they could replace jejunal biopsy in patients with a high pretest probability of CD.

Patients and Methods: This retrospective study included sera from 149 CD patients and 119 controls, all with intestinal biopsy. All samples were analysed for IgA and IgG antibodies against native gliadin and deamidated gliadin peptides (dpgli) and for IgA antibodies against tissue transglutaminase and endomysium.

Results: dpgli for IgG antibody determination were superior to native gliadin: 68% vs. 92% specificity and 79% vs. 85% sensitivity for dpgli and native gliadin, respectively. Positive (76% vs. 93%) and negative (72% vs. 83%) predictive values were also higher for dpgli compared to native gliadin. For IgA gliadin antibody determination, sensitivity improved from 61% to 78% with dpgli, whereas specificity and positive predictive value remained at 97% ($p<0.00001$). A combination of four tests (IgA and IgG anti-dpgli and IgA anti-tissue transglutaminase and anti-endomysium) yielded positive and negative predictive values of 99% and 100%, respectively. Omitting the endomysium antibody determination still yielded positive and negative predictive values of 99% and 98%, respectively.

Conclusion: dpgli yielded superior results compared to native gliadin. A combination of three or four antibody tests including IgA anti-tissue transglutaminase and/or anti-endomysium permitted diagnosis or exclusion of CD without intestinal biopsy in a high proportion of patients (78%). Jejunal biopsy would be necessary in patients with discordant antibody results (22%). With this two-step procedure, only patients with no CD-specific antibodies would be missed.
INTRODUCTION

Celiac disease (CD) is an immune-mediated enteropathy that is caused by intolerance to gluten in genetically susceptible individuals. The prevalence among the European population is approximately 1%;[1, 2] it is even higher in the elderly [3]. Thus, CD is one of the most frequent lifelong diseases.

Serological tests for the diagnosis of CD have improved substantially in the last 20 years. In 1998,[4] we proposed a low-risk and cost-effective algorithm for diagnosing various forms of gluten-sensitive enteropathy with a positive predictive value (ppv) of 99%, using a combination of different antibody determinations, namely anti-endomysium (EMA), IgA anti-tissue transglutaminase (IgA anti-tTG), and IgA and IgG anti–native gliadin (IgA and IgG anti-ngli). In a population with a high pretest probability of disease, synchronous determination of three or four CD-specific antibodies has a very high ppv and negative predictive value (npv) and can eliminate the necessity of small-bowel biopsy in many patients suspected of having CD.[4]

In recent years, the use of ngli as an antigen in antibody-detection tests was replaced by deamidated gliadin peptides (dpgli), which perform better diagnostically than ngli.[5–12] Our goal in this study was to investigate whether using dpgli instead of ngli, alone or in combination with other tests (EMA and IgA anti-tTG), reduces the number of jejunal biopsies without missing CD patients during diagnostic procedure.

PATIENTS AND METHODS

Included in this retrospective study were serum samples from 268 patients who underwent biopsies because of clinically suspected CD. The patients suffered from gastrointestinal symptoms such as diarrhoea, constipation, poor weight gain, chronic vomiting, abdominal pain, flatulence, and failure to thrive, or disorders such as unexplained weight loss in adults, iron-deficiency anaemia, lassitude, psychic
disorders, short stature, and diabetes type I. IgA-deficient CD patients were
excluded. The histology of 149 patients (104 females, age range at biopsy 0.9–80
years, median 29 years; 45 males, age range at biopsy 2–73 years, median 13
years) revealed subtotal or total villous atrophy, hyperplasia of the crypts, and an
increase in intraepithelial lymphocytes (Marsh classification 3a, b, or c lesions). All of
these patients recovered after starting a gliadin-free diet and were regarded as
having CD. The biopsies of 119 patients (66 females, age range at biopsy 1.5–72
years, median 17 years; 53 males, age range at biopsy 2–66 years, median 7 years)
revealed normal mucosa or mucosa with slight, nonspecific changes; these patients
were considered to be free of CD. All of the patients were on gluten-containing diets
at the time of blood sampling. Serum for antibody determinations was obtained within
2 months prior to endoscopic intervention.

All of the samples were analysed for antibodies against tTG, ngli, and dpgli by fully
automated FEIA tests (Elia Celikey IgA, Elia Gliadin IgA, Elia Gliadin IgG, Elia Gliadin
DP IgA and Elia Gliadin DP IgG; Phadia [now Thermo Fisher Scientific],
Freiburg, Germany) performed on the Phadia 100 instrument. For evaluation of the
results, we did not use the cut-off values indicated by the manufacturer. Instead, for
each antigen, we determined the best cut-off value within our sample in terms of the
sum of false-positive and false-negative results: IgA anti tTG = 7, IgA anti-ngli = 7,
IgG anti-ngli = 7, IgA anti-dpgli = 7, and IgG anti-dpgli = 10. EMA was analysed by
indirect immunofluorescence on monkey oesophagus sections: cut-off = serum
dilution: 1:5.
STATISTICAL TOOLS

Beside the usual descriptors for diagnostic tests, such as the sensitivity, the specificity, the positive and negative predictive values, the quantity „efficiency“ was used: the efficiency of a diagnostic test is its percentage of correct outcomes. Contingency tables were evaluated with “Fisher’s Exact Test”. To test whether the replacing or adding of a diagnostic test improves the outcome at a statistical significant degree, “McNemar’s Test for Significant Changes” was applied.

Both tests were used in their precise form (not only asymptotic) as available in the software package StatXact version 6.3.0. from Cytel Software Corporation Cambridge, MA, U.S.A.

RESULTS

Deamidated gliadin peptides compared with native gliadins as antigens

Sera from 149 patients with CD and 119 control patients were tested for IgG and IgA antibodies against dpgli and ngli proteins. IgG antibody determination for dpgli was superior to ngli; the specificity was 68% vs. 92% and sensitivity was 79% vs. 85% for ngli and dpgli, respectively, while ppv was 76% vs. 93% and npv was 72% vs. 83% for ngli and dpgli, respectively. For IgA antibody determination, the sensitivity was 61% vs 78% for ngli and dpgli, respectively, while the specificity and ppv remained at a high level of 97% (McNemar’s test for significant changes $p < 0.00001$, Table 1). Because dpgli antigens were clearly superior to ngli, we only used anti-dpgli for further CD-specific antibody determinations.

Antibody profile in CD and control patients

We also determined the levels of IgA anti-tTG and EMA in sera from the 149 CD patients and 119 controls (Table 2). Because the IgA anti-tTG and EMA results were
comparable, we have omitted showing EMA; instead, we have shown the IgA anti-tTG and IgA and IgG anti-dpgli antibody levels of each individual and compared them with the histological result. We used a multiple test consisting of three individual tests that produces eight possible results. We defined the outcome of the multiple tests as positive only when all three individual tests were above the cut-off, and as negative only when all three individual tests were below the cut-off. The majority of the patients (208/268) had either positive (110) or negative (98) results in all three tests. Nearly all patients (109/110) positive for antibodies in all three tests had CD according to histological findings. The ppv was 99% in our population with 59% frequency of CD. Patients who did not test positive for CD-specific antibodies in any of the three tests were almost all free of celiac disease according to the results of jejunal biopsy (96/98 patients); the npv was 98%. Patients with discordant antibody results (60/268 patients, 22%) could not be defined as positive or negative for CD with the multiple tests and remained unclassified. This indicates that a biopsy is avoidable if all antibody values are either above or below the cut-off. In patients with discordant antibody results, an intestinal biopsy is necessary to diagnose or exclude CD.

**Performance of single antibody tests and selected test combinations**

We compared the performance of IgA and IgG anti-dpgli, IgA anti-tTg, and EMA tests and calculated the sensitivity, specificity, ppv, npv and efficiency of each test and some of the possible test combinations (Table 3). We also indicated the absolute number of patients whose antibody test results were false positive or false negative for CD, as well as those who could not be classified based on antibody tests. Most of the following diagnostic tests are multiple tests (compare to Table 2). We defined the outcome of a multiple test as positive only when all individual tests were above the cut-off, and as negative only when all individual tests were below the cut-off. Test
combinations containing only IgA antibodies were not considered; they are not suitable for diagnostic purposes, because of the possibility that some patients may be deficient in IgA.

Currently, biopsies are often performed when IgA anti-tTG or EMA is found. Negative serological results are usually not followed by a jejunal intervention except when there is a very strong clinical suspicion of CD. However, the data in Table 3 clearly show that single tests are neither specific nor sensitive enough to reduce the number of biopsies in patients with symptoms of CD, although the number of nonclassified patients was zero. Single tests such as the widely used IgA anti-tTG test can give rise to many falsely classified patients.

A combination of two tests also yielded many incorrectly classified patients and is therefore not suitable for reducing the number of biopsies. The two-test combinations yielded either too many false-positive diagnoses (IgG anti-dpgli + IgA anti-tTG or IgG anti-dpgli + EMA) or too many false-negative diagnoses (IgA anti-dpgli + IgG anti-dpgli), although the number of nonclassified patients was smaller than in combinations with more than two tests. The combination of four tests was optimal, as only one patient was false positive, no patients were false negative, the ppv was 99%, and the npv was 100%. For practical reasons, a combination of three tests using IgA anti-tTG instead of EMA in combination with IgA and IgG anti-dpgli (shown in Table 2) may be sufficient to set a standard (ppv 99%, one false-positive result; npv 98%, two false-negative results). In 208/268 patients (78%), a biopsy could be avoided, while 60/268 (22%) could not be diagnosed with serological tests because their results were in disagreement (only one or two results above the cut-off, with the remaining result below the cut-off, table 2).
DISCUSSION

The diagnosis of CD has traditionally depended upon intestinal biopsies and has been extended to include an array of serological markers. The guidelines from the European and North American societies for gastroenterology always require a biopsy for diagnosis.[13,14] CD is diagnosed when the duodenal and jejunal mucosa display villous atrophy, crypt hyperplasia, and an increase in intraepithelial lymphocytes.[15–17] However, different diseases not related to gluten-sensitive enteropathy can induce a flat mucosa, thus mimicking CD. Moreover, patients with gluten-sensitive enteropathy and normal small bowel mucosal architecture have also been described.[18-22] Most likely because of a lack of technical proficiency, either with grasping biopsy forceps or endoscopic procedure, biopsy specimens have been shown to be sufficient for diagnosis of CD in only 90% of cases.[23] Furthermore, CD may be missed during histological examinations due to variations in the assessments of different pathologists.[24] Because of this, and because of the inconvenience and high cost associated with jejunal biopsy and the high prevalence of CD in the general population, less-invasive tests are required. In the last 20 years, serological tests for the diagnosis of CD have improved substantially.[25–31] The data contained in table 3 indicate, however, that the criteria for choosing the best tests must be defined. The best test for clinicians who want to reduce the number of jejunal interventions in a population with a high frequency of CD is the one with the lowest sum of false-positive and false-negative diagnoses; that is, the test with the highest ppv and the highest npv. In our study, a combination of four antibody tests yielded a ppv of 99% and an npv of 100%. Multiple tests result in high predictive values, but leave many patients with discordant antibody results, who must then undergo jejunal biopsy as the second step in diagnosis. For practical reasons, we may omit EMA in our combination of antibody tests, and instead chose the test combination of IgG anti-
dpgli + IgA anti-dpgli + IgA anti-tTG (Tables 2 and 3), with a ppv of 99% and an npv of 98%, as the first step in our diagnostic procedure. Out of 268 patients, 208 (78%) were correctly classified with the serological tests; i.e., they had either three tests above or three tests below the cut-off. The second step, intestinal biopsy, was necessary in the remaining 60 patients (22%), who had discordant antibody results. This two-step diagnostic procedure reduces the number of intestinal biopsies and increases the sensitivity of the entire diagnostic procedure; only CD patients without any CD-specific antibodies would be missed.

In 1998,[4] we suggested the above combination of serological tests as a low-risk and cost-effective algorithm for diagnosing various forms of CD; this combination—still using anti-ngl—was confirmed [31,32] in a total of 1,873 patients with jejunal biopsy. Because of preselection according to symptoms, the prevalence of CD was 59%. The ppv of three tests with congruent positive antibody results was >99%. The npv of all three antibody tests was 98%. However, 37% (599/1,873) of the patients with discordant antibody results could not be classified by antibody tests alone. In the present study, the number of nonclassified patients was reduced from 37% to 22% ($p < 0.001$), due to the better performance of anti-dpgli compared to anti-ngli. Thus, in a population with a high pretest probability of CD, using a combination of three or four antibody tests should obviate the need for as many as 78% of jejunal biopsies.

Recently, several studies have questioned the necessity of performing a jejunal biopsy on all individuals with suspected CD.[33–37] One approach defined five criteria, including clinical signs, of which four had to be fullfilled for a diagnosis of CD.[36] Other studies describe the association of very high IgA anti-tTG antibody-titers with Marsh 3 histopathology.[33,34,35] Recent guidelines, released by the European Society for Pediatric Gastroenterology, Hepathology and Nutrition stated that in patients with high anti-tTG antibody titers (>10 times the upper limit of normal)
intestinal biopsy is redundant.[38] These proposals attain a high specificity and will result in few patients with a false positive diagnosis. However, many CD patients do not have a such high anti-TTG titers; therefore, they will require intestinal biopsy. Sugai et al.[37] investigated single antibody tests and various combinations of two antibody tests in populations with different pretest probabilities for CD. As in our study, sensitivity was lower in combined tests than in single tests, but ppv increased significantly not only in the population with high but also in the group with low pretest probability for CD. They concluded that “appropriate use of CD serology might accurately identify the vast majority of CD patients in populations with different pretest probabilities.”

CONCLUSION
There is no single test—not even jejunal biopsy—that can conclusively diagnose or exclude CD in every individual. Therefore, we propose the following two-step diagnostic procedure: 1st step: Combined, simultaneous determination of IgA and IgG anti-dpgli + IgA anti-tTg and/or EMA. The vast majority of patients will have either three positive or three negative results, obviating the need for a biopsy. 2nd step: Jejunal biopsy should be performed only in patients with discordant antibody results; i.e., in patients who cannot be classified as having CD by antibody tests alone. In any case, the clinical effects of a gluten-free diet must be controlled.

Funding: none declared
Conflict of interests: none declared
Authors' contributions
Study design ABW,FH; Patients selection: ABW,FH; Sample analysis, interpretation ABW,FH, Statistics: MB, Manuscript writing: ABW,MB,FH
All authors read and approved the final manuscript.
References


13. Report of a working group of the united european gastroenterolgy week in Amsterdam. When is a coeliac a coeliac?. Eur J Gastroenterol Hepatol 2001; 13:1123-1128

34. Donaldson MR, Book LS, Leiferman KM, Zone JJ, Neuhausen SL. Strongly positive tissue transglutaminase antibodies are associated with Marsh 3 histopathology in adult and pediatric celiac disease.J Clin Gastroenterol. 2008 Mar;42(3):256-60
### Table 1.

Anti-dpgli antibody tests vs. anti-ngli antibody tests in 149 CD patients and 119 controls.

<table>
<thead>
<tr>
<th></th>
<th>IgA anti-ngli</th>
<th>IgA anti-dpgli</th>
<th>IgG anti-ngli</th>
<th>IgG anti-dpgli</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>61%</td>
<td>78%</td>
<td>79%</td>
<td>85%</td>
</tr>
<tr>
<td>Specificity</td>
<td>97%</td>
<td>97%</td>
<td>68%</td>
<td>92%</td>
</tr>
<tr>
<td>Positive predictive value</td>
<td>97%</td>
<td>97%</td>
<td>76%</td>
<td>93%</td>
</tr>
<tr>
<td>Negative predictive value</td>
<td>67%</td>
<td>78%</td>
<td>72%</td>
<td>83%</td>
</tr>
</tbody>
</table>

McNemar's test for significant changes $p < 0.00001$.

61 correct changes; 154 remain correct; 7 wrong changes; 46 remain wrong;
<table>
<thead>
<tr>
<th>IgA</th>
<th>IgA</th>
<th>IgG</th>
<th>CD</th>
<th>Controls</th>
<th>Total</th>
<th>Classification</th>
<th>Predictive value</th>
</tr>
</thead>
<tbody>
<tr>
<td>anti-tTg</td>
<td>anti-dpgli</td>
<td>anti-dpgli</td>
<td>n = 149</td>
<td>n = 119</td>
<td>n = 268</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+</td>
<td>+</td>
<td>+</td>
<td>109</td>
<td>1</td>
<td>110</td>
<td>110 positives</td>
<td>99%</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
<td>-</td>
<td>7</td>
<td>1</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+</td>
<td>-</td>
<td>+</td>
<td>15</td>
<td>4</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+</td>
<td>-</td>
<td>-</td>
<td>13</td>
<td>10</td>
<td>23</td>
<td>60 not</td>
<td></td>
</tr>
<tr>
<td>-</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>classified</td>
<td></td>
</tr>
<tr>
<td>-</td>
<td>+</td>
<td>-</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>+</td>
<td>3</td>
<td>5</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>96</td>
<td>98</td>
<td>98 negatives</td>
<td>2%</td>
</tr>
</tbody>
</table>

+ antibody present; - antibody absent
Table 3.
Performance of single and selected combinations of antibody tests (n = 268, 149 CD patients and 119 controls)

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>fp</th>
<th>nc</th>
<th>sens</th>
<th>spec</th>
<th>ppv</th>
<th>npv</th>
<th>effic</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>single tests</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IgA anti-dpgli</td>
<td>33</td>
<td>4</td>
<td>0</td>
<td>78%</td>
<td>97%</td>
<td>78%</td>
<td>86%</td>
<td></td>
</tr>
<tr>
<td>IgG anti-dpgli</td>
<td>22</td>
<td>10</td>
<td>0</td>
<td>85%</td>
<td>92%</td>
<td>83%</td>
<td>88%</td>
<td></td>
</tr>
<tr>
<td>IgA anti-tTG</td>
<td>5</td>
<td>16</td>
<td>0</td>
<td>97%</td>
<td>87%</td>
<td>90%</td>
<td>95%</td>
<td></td>
</tr>
<tr>
<td>EMA</td>
<td>3</td>
<td>18</td>
<td>0</td>
<td>98%</td>
<td>85%</td>
<td>89%</td>
<td>97%</td>
<td></td>
</tr>
<tr>
<td><strong>Combinations of two tests</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IgG anti-dpgli + IgA anti-tTG</td>
<td>2</td>
<td>5</td>
<td>39</td>
<td>83%</td>
<td>82%</td>
<td>96%</td>
<td>83%</td>
<td></td>
</tr>
<tr>
<td>IgG anti-dpgli + EMA</td>
<td>1</td>
<td>6</td>
<td>39</td>
<td>84%</td>
<td>82%</td>
<td>95%</td>
<td>99%</td>
<td></td>
</tr>
<tr>
<td>IgA anti-dpgli + IgG anti-dpgli</td>
<td>15</td>
<td>1</td>
<td>37</td>
<td>73%</td>
<td>89%</td>
<td>88%</td>
<td>80%</td>
<td></td>
</tr>
<tr>
<td><strong>Combinations of three tests</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IgA anti-dpgli + IgG anti-dpgli + EMA</td>
<td>1</td>
<td>1</td>
<td>62</td>
<td>72%</td>
<td>81%</td>
<td>99%</td>
<td>99%</td>
<td></td>
</tr>
<tr>
<td>IgA anti-dpgli + IgG anti-dpgli + IgA anti-tTG*</td>
<td>2</td>
<td>1</td>
<td>60</td>
<td>73%</td>
<td>81%</td>
<td>99%</td>
<td>98%</td>
<td></td>
</tr>
<tr>
<td>IgG anti-dpgli + EMA + IgA anti-tTG</td>
<td>0</td>
<td>5</td>
<td>45</td>
<td>83%</td>
<td>80%</td>
<td>96%</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td><strong>Combination of all four tests</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IgG anti-dpgli + IgA anti-dpgli + EMA + IgA anti-tTG</td>
<td>0</td>
<td>1</td>
<td>65</td>
<td>72%</td>
<td>79%</td>
<td>99%</td>
<td>100%</td>
<td></td>
</tr>
</tbody>
</table>

*This test combination is identical to the test in Table 2.
Positive, above the cut-off in all tests; negative, below the cut-off in all tests; fp, number of false-positive patients; fn, number of false-negative patients; nc, number of patients not classified because of discordant antibody results.