Serologic Testing for Celiac Disease in the United States: Results of a Multilaboratory Comparison Study

JOSEPH A. MURRAY,[1]* JUDITH HERLEIN,[2] FRANK MITROS,[2] AND JAMES A. GOEKEN[2]

Departments of Internal Medicine[1] and Pathology,[2] University of Iowa College of Medicine, and the Department of Veterans Affairs Medical Center, Iowa City, Iowa

The aim of this study was to compare the efficiencies of six reference laboratories for serologic testing for celiac disease. Serum from 20 patients with untreated celiac disease and from 20 controls was thawed, divided, and distributed to each participating laboratory, which performed endomysial antibody tests. Five laboratories also performed antigliadin antibody tests. Sensitivity for endomysial antibody immunoglobulin A (IgA) varied from 57 to 90%. In all laboratories, the specificity for celiac disease was 100%. The sensitivity and specificity for both IgA and IgG antigliadin antibody varied significantly. When results from all three tests were combined in each laboratory, sensitivity was 90 to 100%. The specificity for endomysial antibody was 100% in the laboratories. Sensitivity was less than reported previously. Standardization of these tests is needed in the United States.

Celiac disease is a permanent intolerance to gluten that results in damage to the mucosa of the small intestine. This damage consists of mucosal inflammation and loss of absorptive surface area and is manifested by a broad spectrum of symptoms and nutritional deficiencies (7, 15, 23). For almost 30 years, intestinal biopsy has been the standard for the diagnosis of this disease. Although the mucosal damage is primarily cellular, untreated celiac disease is also associated with a humoral immune response that consists of both secreted intestinal and circulating serologic antibodies (14, 20) directed against the reticulin and endomysium of connective tissue, "endomysial antibodies" (EMAs), and against various peptides derived predominantly from wheat, "antigliadin antibodies" (AGAs). EMAs have been proposed as the most reliable serologic marker for celiac disease (8, 25).

Most studies that have examined the usefulness and accuracy of these tests were performed in European laboratories (1, 2, 9–13, 16, 18, 20, 26). Many large studies may be subject to a positive selection bias because of the wide use of serologic tests to refer patients for subsequent biopsy to confirm the diagnosis of celiac disease. The selection bias incurred would overestimate the sensitivity and specificity of the serologic testing in a subsequent retrospective analysis using stored sera from these subjects.

EMAs. The EMA test is an indirect immunofluorescence assay that uses monkey smooth muscle esophagus as a substrate (Fig. 1). Many variables may affect the test, including the light source, level of ambient light, training and experience of the operator, substrate used, and the initial screening dilution. Published results suggest that the endomysial immunoglobulin A (IgA) indirect immunofluorescence assay is the most accurate test available, with a reported sensitivity of 95 to 100% and a specificity of 99 to 100%.

In the United States, serologic testing for celiac disease is performed largely by commercial reference laboratories and laboratories in a few academic institutions, including our own. It is not known how reliable these tests are in the clinical setting in the United States. The aims of our study were (i) to

compare the proficiency of EMA-IgA, AGA-IgG, and AGA-IgA antibody tests in detecting celiac disease in a U.S. population that was not preselected by serologic results and (ii) to identify the variability in testing results among laboratories by using a control group with biopsy-proven celiac disease.

MATERIALS AND METHODS

Residual sera from patients and controls were stored at −70°C and thawed once for the study. Eight reference laboratories known to provide EMA testing were invited to participate, and two declined. The organizing institution (University of Iowa) was excluded from the comparison study to avoid any apparent conflict of interest. The six participating laboratories were Mayo Medical Laboratories (Rochester, Minn.), IMMCO Diagnostics (Buffalo, N.Y.), MRL References Laboratories (Cypress, Calif.), Specialty Laboratories (Santa Monica, Calif.), University of Maryland (Baltimore, Md.), and ARUP (Salt Lake City, Utah).

Each laboratory received a small aliquot of each serum specimen from all 40 subjects. The aliquots were shipped at 4°C. The laboratories were asked to perform the endomysial IgA immunofluorescence assay according to their usual methods. Five of the six laboratories also performed AGA-IgG and AGA-IgA enzyme-linked immunosorbent assay tests by their usual methods. The studies were performed in a single-blind manner by each laboratory.

Patient groups. The 40 subjects (20 patients and 20 controls) had been evaluated for gastrointestinal symptoms, mainly diarrhea. The evaluation included at least three endoscopic duodenal biopsy samples taken from the second, or more distal, portion of the duodenum. All biopsy samples were evaluated by an experienced gastrointestinal pathologist (F.M.). Either serologic tests were not performed before biopsy or the results were not available at the time of biopsy. For the 20 patients, biopsy samples from the small bowel had a histologic pattern typical of celiac disease, with subtotal villous atrophy, and all the patients had a complete response (histologic or clinical) after the institution of a gluten-free diet. The patients with celiac disease were monitored at 6 and 12 months, and 18 of the 20 patients underwent repeat biopsy, which demonstrated complete resolution of intestinal damage. The two patients who did not have a repeat biopsy had complete resolution of their malabsorptive symptoms. The biopsy samples from all the controls were histologically normal. At the time that the biopsy and blood samples were obtained, all patients were following a normal gluten-containing diet. Of the 20 patients with celiac disease (mean age, 50 years; range, 3 to 80 years), 18 were adults.

This study was approved by the Human Subjects Review Board of the University of Iowa.

Statistical analysis. Contingency tables were generated from the results from each laboratory. Sensitivity and specificity between laboratories were compared by χ^2 analysis. The κ coefficient was calculated to examine the agreement among laboratories.

RESULTS

Overall comparison of laboratory results. The overall κ coefficient for agreement among laboratories on all tests, controlled for biopsy, was 0.842 (95% confidence interval, 0.779 to

---

* Corresponding author. Present address: Division of Gastroenterology and Hepatology and Internal Medicine, Mayo Clinic, 200 First St. SW, Rochester, MN 55905.
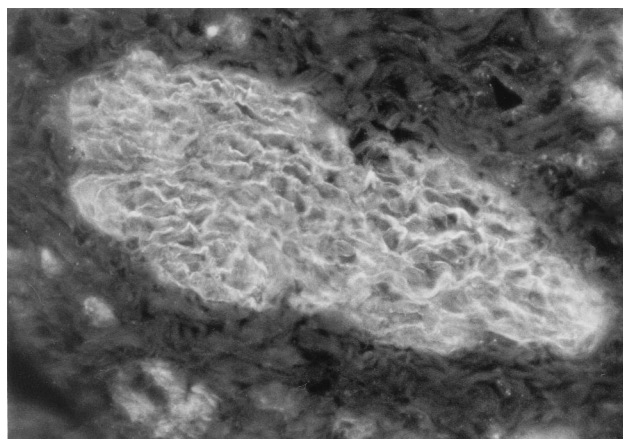
FIG. 1. Endomysial antibody staining of the connective tissue around smooth muscle bundles.
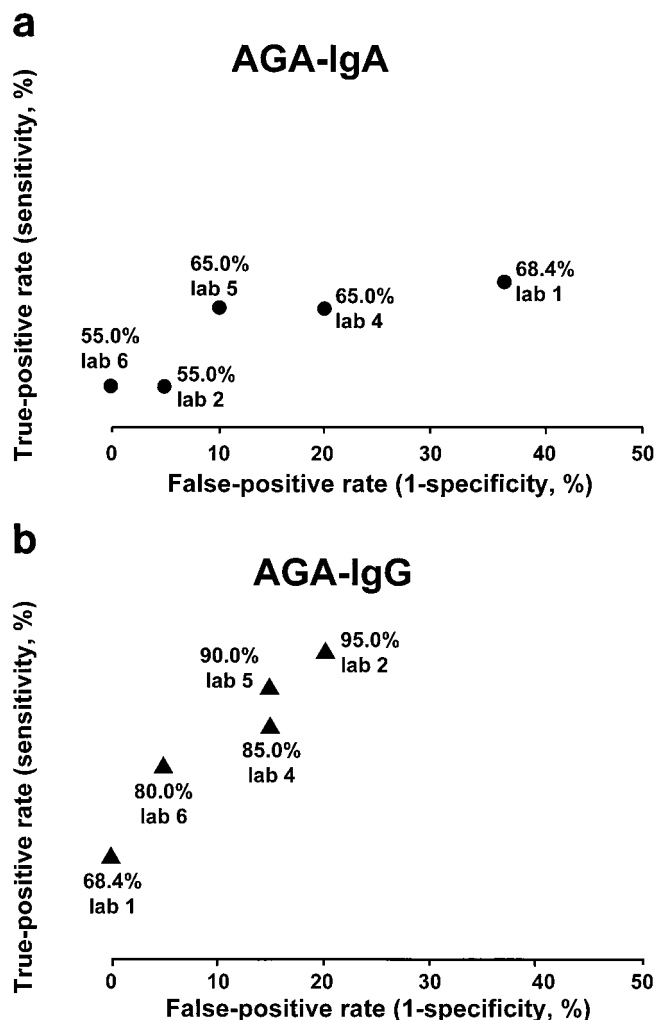

FIG. 2. Receiver operating characteristic showing variability among laboratories and the two tests in the (true-positive) sensitivity versus false-positive (1 − specificity) for gliadin IgA (a) and IgG (b).

0.907). The test for equal κ coefficient showed that one laboratory differed from the others, with a lower degree of agreement with the biopsy-proven diagnosis ($P < 0.03$). This difference was accounted for by the lower sensitivity of the EMA testing for celiac disease.

EMAs. The method used for EMA testing was broadly similar among laboratories. Each laboratory used monkey esophagus as the substrate for the indirect immunofluorescence assay. The serum dilution used for screening varied from 1:20 to 1:2. The primary interpreter of the test varied, from a technologist to a pathologist. Each laboratory required the demonstration of the staining pattern typical of connective tissue surrounding smooth muscle bundles. The specificity of the EMA-IgA test was 100% in all laboratories. The mean sensitivity was 75% (range, 57 to 90%).

The κ coefficient for EMA-IgA tests was 0.739 (95% confidence interval, 0.639 to 0.838), with no significant difference among laboratories. Comparison of titers was not possible because of the different dilution strategies used in the laboratories, and the low volume of serum samples precluded serial dilutions in most laboratories.

Gliadin antibodies. Five of the six laboratories performed AGA-IgA and AGA-IgG antibody tests on the samples. Each laboratory used a different assay system and method for the AGA tests. The reference ranges and units also varied, making quantitative comparison impossible. The specificity and sensitivity of both the AGA-IgA and AGA-IgG tests varied among laboratories (Table 1).

The AGA-IgA tests had the greatest variability and the poorest agreement when controlled for the actual diagnosis and among laboratories (Fig. 2). Two of the 20 patients with celiac disease had selective IgA deficiency. The IgA-based tests were negative for these two patients in all laboratories. One laboratory measured total IgA and detected EMA-IgG in sera

from the two IgA-deficient patients. AGA-IgG tests were positive for the two IgA-deficient patients.

Of the three tests for celiac disease, the overall sensitivity of any positive test of the three done did not differ among laboratories; however, the false-positive rate varied significantly (Fig. 3). This was primarily due to the low specificity of the gliadin IgG and IgA tests in most laboratories.

DISCUSSION

The data from this study confirm the high degree of specificity of the EMA test for celiac disease. It could be argued that a larger control group is needed to test the positive predictive value in a population with a low prevalence of celiac disease. If this degree of positive predictive power is confirmed in broader clinical use, biopsy of the small intestine may not be needed to diagnose celiac disease when the clinical presentation suggests the disease and the EMA test is positive (25). It should be noted that our patients were mainly adults, and it has been suggested that the specificity of this test is lower in children (4).

The coefficient of agreement among the laboratories was high for the EMA test, but the sensitivity varied, with one laboratory having a result statistically different from that of all

TABLE 1. Variation of sensitivity and specificity of AGA and EMA testing among laboratories

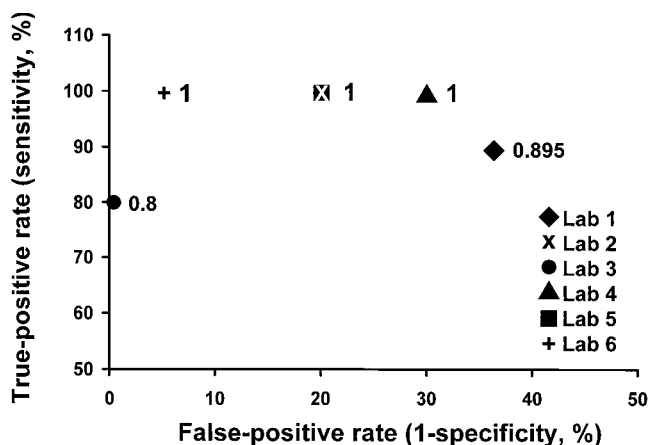| Antibody test	Sensitivity (%)	Specificity (%)
AGA-IgA	55–68.4	63.5–100
AGA-IgG	68.4–95	80–100
EMA-IgA	57.9–90	100

FIG. 3. Receiver operating characteristic using "any positive test," showing an increased false-positive rate when gliadin IgA and IgG results are combined with those of endomysial IgA antibodies to maximize the sensitivity of the screening method. The numbers on the data points correspond to the true positive, where 1.0 is 100%.

the others. The laboratory with the different result used the greatest serum dilution for the initial screen, 1:20. The initial screening dilution may affect both the sensitivity and specificity of the test (21). Our data suggest that a negative EMA-IgA test alone is insufficient to rule out the diagnosis of celiac disease. Our 20 patients with celiac disease all had subtotal villous atrophy, a situation in which the EMA test is thought to be most sensitive; it may be less sensitive in patients with lesser degrees of mucosal damage (19, 24). These results in general are lower than those reported in the literature. Why this is so cannot be explained from our data. The single laboratory that had the lowest sensitivity used a higher dilution for the screening dilution. Other factors such as interpreter variability, buffer conjugates used, and substrate preparation may all affect the accuracy of the tests.

The sensitivity and specificity of AGA tests are known to vary. Our study confirms the results of previous studies that indicated that the specificity of AGA-IgA and AGA-IgG tests does not approach that of the EMA test. Unexpectedly, AGA-IgG tests were more sensitive and more specific than AGA-IgA tests. This difference may be due partly to the two patients with IgA deficiency, but this would not explain the poor specificity. Extrapolating these results to a population with a much lower prevalence of celiac disease (for example, 0.5%), the false-positive tests would greatly outnumber the true-positive tests. Larger studies are under way to examine this question. A positive AGA-IgA test does not replace the need for a more accurate test to make the diagnosis of celiac disease.

The inclusion of the two patients with IgA deficiency did contribute to the decreased sensitivity of the IgA-based tests in all laboratories. However, even after these were excluded, the sensitivity was less than 95% in most laboratories. Screening for IgA deficiency has been suggested as part of the evaluation for celiac disease (5). The AGA-IgG test was positive for two patients who were IgA deficient. These patients were also EMA-IgG positive. The estimated prevalence of selective IgA deficiency in celiac disease varies from 1 to 5% (3, 15); however, if serologic tests are used to detect most cases, then the true prevalence of IgA deficiency in celiac disease may be underestimated. Of unselected persons with selective IgA deficiency, 7.7% had celiac disease (17). It may be prudent to consider including a rapid test for IgA deficiency for patients

who are either completely negative for AGA and EMA or are positive only for AGA-IgG. The lack of AGA-IgG in an IgA-deficient patient does not rule out the possibility of celiac disease (5, 17). It seems necessary to combine AGA-IgA, AGA-IgG, and EMA-IgA testing to maximize sensitivity (Fig. 3). This may be especially important when the suspicion of celiac disease is high; however, when clinical suspicion is very low, there likely will be an unacceptable level of false-positive AGA-IgA and AGA-IgG tests. However, if all three tests are negative, the likelihood of celiac disease is reduced further. In many situations, intestinal biopsies will continue to be required to diagnose celiac disease. Certainly, in patients with symptoms suggestive of malabsorption, intestinal biopsy is mandated not only to diagnose celiac disease but also to identify other mucosal diseases that could result in malabsorption.

The degree of correlation among the laboratory results was unexpected, considering the complete lack of uniformity of testing methods and standards and in the training of the persons who interpreted the EMA tests. This illustrates the robust specificity of this test in clinical practice. The methods for the AGA tests need to be standardized, as do the assay units and reference ranges. Such standards are desirable to both laboratories and clinicians requesting the tests. Standardization will reduce confusion about the interpretation and the use of these tests in clinical practice.

Although commercial contracting for specialized testing has cost advantages and decreases the need for specialized training in many laboratories, it puts distance between the physician and the laboratory performing the test. This makes clinical feedback almost impossible and reduces the likelihood of ongoing validation of the tests. Serologic testing for celiac disease should be included in the College of American Pathologists validation system. The European Union has sponsored an ongoing effort to standardize testing and has an agreed-upon methodology for EMA testing and basic requirements for AGA testing (22). A similar effort is needed in North America. The lack of standardization of these tests, especially the AGA tests, will likely restrict their clinical usefulness.

It has been suggested recently that tissue transglutaminase is the antigen recognized by endomysial staining (6). Enzyme-linked immunosorbent assay tests are being developed to identify antibodies to tissue transglutaminase. However, our experience suggests that these tests (and any others that may be developed) should be validated not only against other serologic tests but also against biopsy-based diagnoses, which include a significant proportion of patients who were not referred for biopsy because of positive findings on serologic testing. Those responsible for performing these tests or for selecting a reference laboratory would be advised to carefully validate the test method. How positive sera are selected may influence the apparent accuracy of the validation process. Using sera that are positive for other serologic tests of celiac disease may not be sufficient to predict the sensitivity of the test. Verifying the serologic test by histologic abnormality would be a valuable check, but correlation usually is not available to the reference laboratory. Often only the requesting physician may be aware of the disparity between serologic and subsequent biopsy (if it is done) results. Recurring disparity will often cause the primary physician to discard the use of the test because of the perception that it is ineffective or inaccurate.

Our observation was that although a combination of all three tests maximized the sensitivity, the specificity of this strategy was low. The clinician faced with a subject in whom the probability of celiac disease is high (>10%) probably should consider an intestinal biopsy. If the pretest likelihood of disease is approximately 5%, it may be reasonable to infer a

strong negative predictive value if all three tests are negative, but if any one test is positive, biopsy is needed. Although this approach may be acceptable for a patient with symptoms, it would not be feasible in a population-based screening project with an unacceptably high rate of normal biopsy findings.

The sera in the present study and in most validation studies of serologic tests were from patients with subtotal villous atrophy. The accuracy of the tests must be studied for patients with less than total villous atrophy, because the results of their tests may differ from those of patients with classic celiac disease as defined by total villous atrophy.

The results of our study support the high specificity of the EMA-IgA test for identifying celiac disease in the United States. However, these results must be confirmed in a larger study group before small-bowel biopsy may be deemed unnecessary for patients with suggestive symptoms and positive EMA test results. In the present study, the sensitivity of the EMA-IgA test was less than that reported previously and varied greatly among laboratories.

ACKNOWLEDGMENTS

REFERENCES

1. **Ascher, H., M. Hahn-Zoric, L. A. Hanson, A. F. Kilander, L. A. Nilsson, and H. Tlaskalova.** 1996. Value of serologic markers for clinical diagnosis and population studies of coeliac disease. Scand. J. Gastroenterol. **31:**61–67.
2. **Bodvarsson, S., I. Jonsdottir, J. Freysdottir, J. N. Leonard, L. Fry, and H. Valdimarsson.** 1993. Dermatitis herpetiformis—an autoimmune disease due to cross-reaction between dietary glutenin and dermal elastin? Scand. J. Immunol. **38:**546–550.
3. **Cataldo, F., V. Marino, G. Bottaro, P. Greco, and A. Ventura.** 1997. Celiac disease and selective immunoglobulin A deficiency. J. Pediatr. **131:**306–308.
4. **Chan, K. N., A. D. Phillips, R. Mirakian, and J. A. Walker-Smith.** 1994. Endomysial antibody screening in children. J. Pediatr. Gastroenterol. Nutr. **18:**316–320.
5. **Dickey, W., S. A. McMillan, E. E. McCrum, and A. E. Evans.** 1997. Association between serum levels of total IgA and IgA class endomysial and antigliadin antibodies: implications for coeliac disease screening. Eur. J. Gastroenterol. Hepatol. **9:**559–562.
6. **Dieterich, W., T. Ehnis, M. Bauer, P. Donner, U. Volta, E. O. Riecken, and D. Schuppan.** 1997. Identification of tissue transglutaminase as the autoantigen of celiac disease. Nat. Med. **3:**797–801.
7. **Ferguson, A.** 1997. Celiac disease, an eminently treatable condition, may be underdiagnosed in the United States. Am. J. Gastroenterol. **92:**1252–1254.
8. **Ferreira, M., S. L. Davies, M. Butler, D. Scott, M. Clark, and P. Kumar.** 1992. Endomysial antibody: is it the best screening test for coeliac disease? Gut **33:**1633–1637.
9. **Friis, S. U., and E. Gudmand-Hoyer.** 1986. Screening for coeliac disease in adults by simultaneous determination of IgA and IgG gliadin antibodies. Scand. J. Gastroenterol. **21:**1058–1062.
10. **Hallstrom, O.** 1989. Comparison of IgA-class reticulin and endomysium antibodies in coeliac disease and dermatitis herpetiformis. Gut **30:**1225–1232.
11. **Kapuscinska, A., T. Zalewski, T. P. Chorzelski, J. Sulej, E. H. Beutner, V. Kumar, and T. Rossi.** 1987. Disease specificity and dynamics of changes in IgA class anti-endomysial antibodies in celiac disease. J. Pediatr. Gastroenterol. Nutr. **6:**529–534.
12. **Kilander, A. F., G. Dotevall, S. P. Fallstrom, R. E. Gillberg, L. A. Nilsson, and A. Tarkowski.** 1983. Evaluation of gliadin antibodies for detection of coeliac disease. Scand. J. Gastroenterol. **18:**377–383.
13. **Lazzari, R., U. Volta, F. B. Bianchi, A. Collina, and E. Pisi.** 1984. R1 reticulin antibodies: markers of celiac disease in children on a normal diet and on gluten challenge. J. Pediatr. Gastroenterol. Nutr. **3:**516–522.
14. **Maki, M., T. Huupponen, K. Holm, and O. Hallstrom.** 1995. Seroconversion of reticulin autoantibodies predicts coeliac disease in insulin dependent diabetes mellitus. Gut **36:**239–242.
15. **Marsh, M. N.** 1992. Gluten, major histocompatibility complex, and the small intestine. A molecular and immunobiologic approach to the spectrum of gluten sensitivity ('celiac sprue'). Gastroenterology **102:**330–354.
16. **McMillan, S. A., D. J. Haughton, J. D. Biggart, J. D. Edgar, K. G. Porter, and T. A. McNeill.** 1991. Predictive value for coeliac disease of antibodies to gliadin, endomysium, and jejunum in patients attending for jejunal biopsy. BMJ **303:**1163–1165.
17. **Meini, A., N. M. Pillan, V. Villanacci, V. Monafo, A. G. Ugazio, and A. Plebani.** 1996. Prevalence and diagnosis of celiac disease in IgA-deficient children. Ann. Allergy Asthma Immunol. **77:**333–336.
18. **Not, T., A. Ventura, S. Peticarari, S. Basile, G. Torre, and D. Dragovic.** 1993. A new, rapid, noninvasive screening test for celiac disease. J. Pediatr. **123:**425–427.
19. **Rossi, T. M., V. Kumar, A. Lerner, L. A. Heitlinger, N. Tucker, and J. Fischer.** 1988. Relationship of endomysial antibodies to jejunal mucosal pathology: specificity towards both symptomatic and asymptomatic celiacs. J. Pediatr. Gastroenterol. Nutr. **7:**858–863.
20. **Seah, P. P., L. Fry, M. A. Rossiter, A. V. Hoffbrand, and E. J. Holloborow.** 1971. Anti-reticulin antibodies in childhood coeliac disease. Lancet **ii:**681–682.
21. **Stern, M., M. Adenauer, C. Keilmann, and I. Wascher.** 1997. Standardization of screening tests, p. 35–46. *In* M. Maki, P. Collin, and J. K. Visakorpi (ed.), Coeliac disease. University of Tampere Press, Tampere, Finland.
22. **Stern, M., M. Teuscher, and T. Wechmann.** 1996. Serological screening for coeliac disease: methodological standards and quality control. Acta Paediatr. Suppl. **412:**49–51.
23. **Trier, J. S.** 1993. Diagnosis and treatment of celiac sprue. Hosp. Pract. **28:**41–44.
24. **Troncone, R., N. Caputo, M. Micillo, L. Maiuri, and V. Poggi.** 1994. Immunologic and intestinal permeability tests as predictors of relapse during gluten challenge in childhood coeliac disease. Scand. J. Gastroenterol. **29:**144–147.
25. **Valdimarsson, T., L. Franzen, E. Grodzinsky, T. Skogh, and M. Strom.** 1996. Is small bowel biopsy necessary in adults with suspected celiac disease and IgA anti-endomysium antibodies? 100% positive predictive value for celiac disease in adults. Dig. Dis. Sci. **41:**83–87.
26. **Volta, U., N. Molinaro, L. de Franceschi, D. Fratangelo, and F. B. Bianchi.** 1995. IgA anti-endomysial antibodies on human umbilical cord tissue for celiac disease screening. Save both money and monkeys. Dig. Dis. Sci. **40:**1902–1905.