Diagnosis of Celiac Disease: Current State of the Evidence

Focus of This Summary

This is a summary of a systematic review evaluating the evidence regarding the comparative accuracy (the balance of sensitivity and specificity) and possible adverse consequences (both direct and indirect) of various methods used to diagnose celiac disease. The systematic review included 60 individual studies and 13 previous systematic reviews published from January 1990 through March 2015. The full report, listing all studies and reviews, is available at *www.effectivehealthcare.ahrq.gov/celiac-disease.* This summary is provided to assist in informed clinical decisionmaking. However, reviews of evidence should not be construed to represent clinical recommendations or guidelines.

Background

Celiac disease is an immune-mediated disorder triggered in genetically susceptible individuals by ingestion of foods that contain gluten. The prevalence of celiac disease in the United States has been estimated at approximately 1 percent, but it appears to be increasing for reasons that are unclear. Risk factors for celiac disease include family history, trisomy 21, Turner syndrome, Williams syndrome, and several autoimmune diseases (including type 1 diabetes).

Various serological, endoscopic, and histological tests are used to diagnose celiac disease, as described in Table 1 on page 2. All methods require that a gluten-containing diet be maintained during the diagnostic process. Additionally, human lymphocyte antigen (HLA) typing may be used to eliminate the diagnosis, as celiac disease is strongly associated with *HLA-DQ2* and *HLA-DQ8*. Endoscopy with duodenal biopsy showing villous atrophy is the current gold standard for diagnosing celiac disease, but the procedure is invasive and accompanied by a risk, albeit small, of abdominal pain, bloating, discomfort, bleeding, or perforation. Thus, identifying noninvasive tests that are accurate and have few or no side effects is important.

Guidelines from the American College of Gastroenterology (ACG)* recommend anti-tissue transglutaminase immunoglobulin A (tTG IgA) as the first-line diagnostic test for patients with celiac disease over the age of 2 years.[1] As IgA deficiency is more common in people with celiac disease than in the general population, measuring total IgA should be considered to determine the need for immunoglobulin G (IgG)-based tests (e.g., deamidated gliadin peptide [DGP] IgG), especially if there is a strong likelihood of celiac disease. Alternatively, IgG tests may be combined with tTG IgA during initial testing in high-likelihood individuals. For children younger than 2 years, the ACG guidelines recommend initial testing with tTG IgA, DGP IgA, and DGP IgG. Duodenal biopsy is recommended when there is a high suspicion of celiac disease, even if serological testing is negative. Video capsule endoscopy (VCE) is only recommended as the initial diagnostic procedure in people unwilling or unable to undergo biopsy.

This systematic review sought to determine the comparative accuracy and possible harms of various methods used to diagnose celiac disease in children or adults. The effectiveness of these tests, both alone and in combination, was determined in various populations of special interest to physicians who treat celiac disease. This review included only studies in which all participants underwent biopsy, regardless of their serological test results. Because a biopsy is invasive, some accuracy studies required that only participants with positive serological tests for celiac disease undergo biopsy. These studies were excluded because the false-negative rate (and thus overall accuracy) of the tests they evaluated could not be determined.

* The ACG guidelines are the only guidelines from the United States referenced in the National Guideline Clearinghouse database at *guideline.gov*.

Conclusions

Current evidence on the accuracy of tests used to diagnose celiac disease supports the excellent sensitivity and specificity of tTG IgA tests (high strength of evidence [SOE]). Endomysial antibody (EmA) IgA tests have lower sensitivity but similar specificity to tTG IgA tests (high SOE). tTG IgA tests are more sensitive than DGP IgA and DGP IgG tests in adults who are not IgA deficient (moderate SOE).

Both tTG and DGP tests are more sensitive in children than in adults. Limited evidence suggests that DGP tests are more accurate than tTG tests in children <24 months of age (low SOE); however, these findings need to be replicated in future studies.

Algorithms combining tTG with either EmA or DGP tests appear to be accurate in both children and adults, although any increased accuracy when they are compared with individual tests is rarely clinically significant. Evidence is insufficient to determine the comparative accuracy of different types of algorithms. The SOE is also insufficient to determine the accuracy of diagnostic tests in specific populations, including those with IgA deficiency, risk factors for celiac disease, or previous negative serological testing for celiac disease.

Testing for the presence of *HLA-DQ2* or *HLA-DQ8* has excellent sensitivity and negative predictive value (high SOE). It is useful for ruling out but not diagnosing celiac disease.

VCE is a safe and accurate means of diagnosing celiac disease in adults who wish to avoid biopsy, although it has a 0.9- to 4.6-percent capsule retention rate. Pooled results reveal that most serological tests have higher sensitivity and specificity than VCE. No data are available on how VCE accuracy varies by population characteristics or setting. Endoscopy with multiple duodenal biopsies has a very low risk of adverse events.

1. Rubio-Tapia A, Hill ID, Kelly CP, et al; American College of Gastroenterology. ACG clinical guidelines: diagnosis and management of celiac disease. Am J Gastroenterol. 2013 May;108(5):656-76. PMID: 23609613.

Agency for Healthcare Research and Quality
Effective Health Care Program

Table 1: Tests Currently Used To Diagnose Celiac Disease

Test	Description	Notes
Anti-tissue transglutaminase, IgA	Tissue transglutaminase (tTG) is an enzyme involved in the cross-linking of certain proteins.	This test is included in the algorithms of all recent guidelines. Other tests may be ordered for individuals who are IgA deficient.
Endomysial antibodies, IgA	The endomysium is the thin connective-tissue layer covering individual muscle fibers. Endomysial antibodies (EmA) develop when the intestinal lining is damaged.	Most patients with active celiac disease and many others with dermatitis herpetiformis have anti-EmA IgA. EmA is included in algorithms of recent diagnostic guidelines but is not as widely used in the United States as in other countries. EmA IgA is less useful in individuals whose IgA is low.
Anti-deamidated gliadin peptide, IgA and IgG	Gliadins are one of the two main groups of proteins in gluten.	Elevated antibodies against deamidated gliadin peptide are often seen in patients who have celiac disease and are on a gluten-containing diet. This is a newer test that may give a positive result in some individuals with celiac disease who are anti-tTG negative, including children <2 years of age.
HLA typing (*HLA-DQ2* or *HLA-DQ8*)	Researchers hypothesize that HLA molecules present gluten antigens to T-cells, which in turn induce tissue damage.	Approximately 95 percent of patients with celiac disease have the *HLA-DQ2* heterodimer, while the remaining 5 percent have the *HLA-DQ8* heterodimer. Lack of these heterodimers all but rules out celiac disease and genetic susceptibility for the disorder.
Video capsule endoscopy (VCE)	VCE involves the ingestion of a capsule containing a tiny camera.	This test is used in individuals (primarily adults) who want to avoid biopsy. It provides high-quality visual evidence of the scalloping, fissuring, and flattened folds associated with celiac disease. One limitation of VCE is that its accuracy depends on the expertise of the individual interpreting the images. Furthermore, the cost of VCE is often not covered by health insurance companies in the United States.
Endoscopy with duodenal biopsy	The histology of a duodenal biopsy is assessed according to the Modified Marsh Classification system (grades 0–3) or qualitatively.	Villous atrophy disclosed by biopsy plus clinical remission while on a gluten-free diet represent the internationally accepted gold standard for celiac disease diagnosis. The clinician should instruct the patient to stay on a gluten-containing diet before biopsy in most circumstances. Marsh grade 1 or grade 2 lesions—in the absence of clinical or serological evidence—are nonspecific and suggestive of celiac disease but not confirmatory of it. Marsh grade 3 lesions are the classic celiac disease lesions and are characteristic but not diagnostic of celiac disease.

HLA = human leukocyte antigen; IgA = immunoglobulin A; IgG = immunoglobulin G

Strength of Evidence Scale*

High:	●●●	High confidence that the evidence reflects the true effect. Further research is very unlikely to change our confidence in the estimate of effect.
Moderate:	●●○	Moderate confidence that the evidence reflects the true effect. Further research may change our confidence in the estimate of effect and may change the estimate.
Low:	●○○	Low confidence that the evidence reflects the true effect. Further research is likely to change our confidence in the estimate of effect and is likely to change the estimate.
Insufficient:	○○○	Evidence either is unavailable or does not permit a conclusion.

* The overall evidence grade was assessed based on the ratings for the following domains: study limitations, directness, consistency, precision, and reporting bias. Other domains that were considered, as appropriate, included dose-response association, plausible confounding, and strength of association (i.e., magnitude of effect). For additional details on the methodology used to assess strength of evidence, please refer to: Owens DK, Lohr KN, Atkins D, et al. AHRQ series paper 5: grading the strength of a body of evidence when comparing medical interventions—Agency for Healthcare Research and Quality and the Effective Health-Care Program. J Clin Epidemiol. 2010 May;63(5):513-23. PMID: 19595577.

Overview of Clinical Research Evidence

Table 2: Summary of Key Findings and Strength of Evidence for the Overall Comparative Accuracy of Diagnostic Tests When Using Endoscopy With Biopsy as the Reference Standard

Please note: In this table, the populations used in the studies were mixed (denoted "M"), children only (denoted "C"), or adults only (denoted "A").

In this review, sensitivity and specificity were considered "good" if at least 70.0 percent, "very good" if from 80.0 to 89.9 percent, and "excellent" if 90.0 percent or greater.

Test	Study Type	No. of Studies (Population)	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	PPV (%) (95% CI)	NPV (%) (95% CI)	Summary	SOE
tTG IgA	**Systematic review**	1 (M)	93.0 (91.2 to 94.5)	96.5 (95.2 to 97.5)	NA	NA	Excellent sensitivity and specificity	●●●
		1 (C)	96.4 (94.3 to 97.9)[a]	97.7 (95.8 to 99.0)[a]				
	Other	16	92.5 (89.7 to 94.6)[b]	97.9 (96.5 to 98.7)[b]	89.4 (88.3 to 90.5)	99.0 (98.8 to 99.1)		
EmA IgA	**Systematic review**	1 (M)	68.0 to 100.0	77.0 to 100.0	NA	NA	Lower sensitivity but equal specificity to tTG IgA tests	●●●
		1 (C)	82.6 to 100.0	98.2 (96.7 to 99.1)				
	Other	7	79.0 (71.0 to 86.0)	99.0 (98.4 to 99.4)	78.9 (71.0 to 85.5)	99.1 (98.6 to 99.5)		
DGP IgA	**Systematic review**	1 (M)	87.8 (85.6 to 89.9)	94.1 (92.5 to 97.5)	NA	NA	Not as accurate as tTG IgA tests in adults and older children	●●●
		1 (C)	80.7 to 95.1	90.7 (87.8 to 93.1)				
	Other	2 (M)	97.0; 96.0	90.7; 96.0	NA	NA		
DGP IgG in children and adults who are not IgA deficient[c]	**Systematic review**	1 (A)	75.4 to 96.7	98.5 to 100.0	NA	NA	Not as sensitive as tTG IgA tests in patients who are not IgA deficient	●●○
		1 (C)	80.1 to 98.6	86.0 to 96.9				
	Other	1 (M)	95.0	99.0	NA	NA		
HLA-DQ2* or *HLA-DQ8	**Other**	2	100.0; 100.0	18.2; 33.3	NA	NA	Excellent sensitivity Useful test to rule out rather than diagnose celiac disease	●●●
Video capsule endoscopy	**Systematic review**	1	89.0 (82.0 to 94.0)	95.0 (89.0 to 99.0)	NA	NA	Very good sensitivity and excellent specificity	●●○

[a] Sensitivity and specificity values are for point-of-care testing.

[b] Sensitivity and specificity values were obtained from pooled analyses after excluding data for threshold levels greater than those used in clinical practice (40 U/mL).

[c] Evidence was insufficient regarding the accuracy of DGP IgG in individuals with IgA deficiency.

95% CI = 95-percent confidence interval; A = adults only; C = children only; DGP = deamidated gliadin peptide; EmA = endomysial antibodies; HLA = human leukocyte antigen; IgA = immunoglobulin A; IgG = immunoglobulin G; M = mixed (adults and children); NA = not applicable; NPV = negative predictive value; PPV = positive predictive value; SOE = strength of evidence; tTG = tissue transglutaminase

Table 3: Summary of Findings and Strength of Evidence for Adverse Events Associated With Invasive Diagnostic Tests

Adverse Events	Test	Type and No. of Studies (N[a])	Systematic Review Findings	Findings of Other Studies	SOE
Direct	**Video capsule endoscopy**	**Systematic review:** 1 (150)[b] **Other:** 3 (1,030)	Capsule retention rate[c]: 1.4%	Capsule retention rate: 0.9% to 4.6%	●●●
	Endoscopy with duodenal biopsy	**Systematic review:** 1 (106)[d]	Infection very rare Bleeding very rare (1.6 cases per 1,000 procedures), unless polyps are removed	NA	●●○

[a] N is the number of studies for systematic reviews and the number of subjects for other studies.

[b] This systematic review represents a review of video capsule endoscopy (VCE) for various suspected small-intestine pathologies, as no systematic reviews contained safety data on VCE when used specifically to diagnose celiac disease.

[c] Capsule retention is typically defined as failure to spontaneously eliminate the capsule from the gastrointestinal (GI) tract within 2 weeks. It is usually asymptomatic but may cause small-bowel obstruction or require surgical removal (especially when strictures are present).

[d] This systematic review represents a review of upper GI endoscopy in general, as no systematic reviews contained safety data on upper GI endoscopy or duodenal biopsy when used specifically to diagnose celiac disease.

NA = not available

Other Findings of the Review

HLA tests

- Based on studies for which sensitivity could be calculated, the ACG estimated the negative predictive value of the *HLA-DQ2/HLA-DQ8* combination test at more than 99 percent in diagnosing celiac disease.

Testing algorithms

- All algorithms studied used tTG tests. The strength of evidence was insufficient to determine the comparative accuracy of the different algorithms. Adding an EmA test to a tTG test resulted in increased specificity, with either no change or a slight decrease in sensitivity. Adding a DGP test to a tTG test resulted in increased sensitivity but decreased specificity. However, the increase in accuracy—when compared with individual tests—was rarely clinically significant because the sensitivity and specificity results varied widely between studies, the study populations were diverse, and the evidence base had high heterogeneity.

Biopsy

- Physician adherence to the duodenal biopsy protocol (4+ specimens) recommended by the American Gastroenterological Association in 2006* decreased as the volume of procedures performed per endoscopy suite increased. Adherence increased as the number of gastroenterologists per endoscopy suite increased (●●○).
- Celiac disease-related histological findings are inaccurately reported more often in community settings when compared with academic settings (●●○).
- Increasing the number and location of biopsy specimens increases diagnostic accuracy in both adult and pediatric populations (●●●).
- A minimum 2-week gluten-containing diet is necessary to induce the intestinal changes needed to diagnose celiac disease in adults via duodenal biopsy (●●○), whereas a 2- to 3-month gluten-containing diet may be necessary to diagnose celiac disease in children via duodenal biopsy (●○○).

Subpopulations

- ***Patients with gastrointestinal (GI) symptoms:*** EmA and tTG tests have very good to excellent sensitivity and specificity in this subpopulation (●●●). Based on 1 previous systematic review, EmA IgA testing has a sensitivity of 90 percent (95% confidence interval [CI]: 80.0% to 95.0%) and specificity of 99 percent (95% CI: 98.0% to 100.0%), and IgA tTG testing has a sensitivity of 89 percent (95% CI: 82.0% to 94.0%) and specificity of 98 percent (95% CI: 95.0% to 99.0%).
- ***Patients without GI symptoms:*** There is insufficient evidence to determine the comparative accuracy of diagnostic tests in this subpopulation (○○○).
- ***Children versus adults:*** tTG and DGP tests may be more sensitive in children than in adults (●○○) based on the findings of 2 systematic reviews in adults and children and 2 other studies. In the other studies, the sensitivity of tTG and DGP was 57 to 96 percent in children and 29 to 85 percent in adults.

* Updated guidelines published by the American College of Gastroenterology (ACG) in 2013 recommend 1 to 2 duodenal bulb specimens and at least 4 specimens from the distal duodenum.

Gaps in Knowledge and Limitations of the Evidence Base

- The comparative effects of different diagnostic methods on various important outcomes such as clinical decisionmaking, adherence to gluten-free diets, quality of life, and symptoms are not known, since only a few studies evaluated these outcomes.
- The evidence was insufficient to determine the comparative accuracy of diagnostic tests in certain subpopulations: patients with differing demographic factors (including race or ethnicity), IgA deficiency, or celiac disease risk factors (including type 1 diabetes or a positive family history) and patients who previously tested negative during serological evaluations for celiac disease.
- Evidence regarding the accuracy of serological tests in the asymptomatic general population is very limited. Because biopsies are invasive, most studies assessing the accuracy of serological tests that use biopsy as the reference standard have been conducted in patients who present for testing because of symptoms.
- Evidence was also insufficient to determine the impact of indirect adverse consequences of diagnostic testing for celiac disease, especially those resulting from misdiagnosis, such as unnecessary lifestyle changes, possible social isolation because of false-positive results, or malabsorption and intestinal damage because of false-negative results.

Ordering Information

For electronic copies of this clinician research summary and the full systematic review, visit *www.effectivehealthcare.ahrq.gov/celiac-disease.*

Source

The information in this summary is based on *Diagnosis of Celiac Disease*, Comparative Effectiveness Review No. 162, prepared by the RAND Southern California Evidence-based Practice Center under Contract No. 290-2012-00006-I for the Agency for Healthcare Research and Quality, January 2016. The review is available at *www.effectivehealthcare.ahrq.gov/celiac-disease.* This summary was prepared by the John M. Eisenberg Center for Clinical Decisions and Communications Science at Baylor College of Medicine, Houston, TX.

AHRQ Pub. No. 15(16)-EHC032-3-EF
July 2016