Evidence Synthesis
Number 144

Screening for Celiac Disease: A Systematic Review for the U.S. Preventive Services Task Force

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
5600 Fishers Lane
Rockville, MD 20857
www.ahrq.gov

Contract No. HHSA-290-2012-00015-I, Task Order No. 4

Prepared by:
Pacific Northwest Evidence-Based Practice Center
Oregon Health & Science University
3181 SW Sam Jackson Park Road
Portland, OR 97239
www.ohsu.edu/epc

Investigators:
Roger Chou, MD
Ian Blazina, MPH
Christina Bougatsos, MPH
Katherine Mackey, MD
Sara Grusing, BA
Shelley Selph, MD, MPH

AHRQ Publication No. 14-05215-EF-1
March 2017
This report is based on research conducted by the Pacific Northwest Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (HHSA-290-2012-00015-I, Task Order No. 4). The findings and conclusions in this document are those of the authors, who are responsible for its contents, and do not necessarily represent the views of AHRQ. Therefore, no statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

The information in this report is intended to help health care decisionmakers—patients and clinicians, health system leaders, and policymakers, among others—make well-informed decisions and thereby improve the quality of health care services. This report is not intended to be a substitute for the application of clinical judgment. Anyone who makes decisions concerning the provision of clinical care should consider this report in the same way as any medical reference and in conjunction with all other pertinent information (i.e., in the context of available resources and circumstances presented by individual patients).

This report may be used, in whole or in part, as the basis for development of clinical practice guidelines and other quality enhancement tools, or as a basis for reimbursement and coverage policies. AHRQ or U.S. Department of Health and Human Services endorsement of such derivative products may not be stated or implied.

None of the investigators has any affiliations or financial involvement that conflicts with the material presented in this report.

Acknowledgements

The authors acknowledge AHRQ Medical Officers Tracy Wolff, MD, MPH, and Karen Lee, MD, MPH, as well as current and former members of the U.S. Preventive Services Task Force who contributed to topic discussions.

Suggested Citation

Structured Abstract

**Background:** Unrecognized celiac disease may have adverse effects on morbidity and mortality.

**Purpose:** To review the evidence on screening for celiac disease in asymptomatic adults, adolescents, and children age 3 years or older for the U.S. Preventive Services Task Force.

**Data Sources:** Ovid MEDLINE, Cochrane Central Register of Controlled Trials, and Cochrane Database of Systematic Reviews (to mid-June 2016).

**Study Selection:** Randomized clinical trials, cohort studies, and case-control studies of screening versus no screening, one screening strategy versus another, treatment versus no treatment, or immediate versus delayed treatment that evaluated clinical outcomes; and studies on diagnostic accuracy of serologic tests for celiac disease.

**Data Extraction:** One investigator abstracted data, a second checked data for accuracy, and two investigators independently assessed study quality using predefined criteria.

**Data Synthesis (Results):** We identified no trials of screening for celiac disease. One recent, good-quality systematic review found serologic tests to be accurate for diagnosing celiac disease, but two studies conducted in asymptomatic populations reported lower sensitivity than in studies not restricted to asymptomatic populations. One fair-quality, small (n=40) Finnish treatment trial of screen-detected, asymptomatic adults with positive serologic findings found initiation of a gluten-free diet associated with small improvement in gastrointestinal symptoms versus no gluten-free diet (<1 point on a 7-point scale) at 1 year, with no differences on most measures of quality of life. No withdrawals due to adverse events occurred during the trial.

**Limitations:** Limited or no evidence for all key questions; limited to English-language studies.

**Conclusions:** More research is needed to understand the effectiveness of screening for and treatment of celiac disease in asymptomatic adults, adolescents, and children; accuracy of screening tests; and optimal screening strategies.
# Table of Contents

## Chapter 1. Introduction
- Purpose and Previous U.S. Preventive Services Task Force Recommendation ........................................... 1
- Condition Definition ............................................................................................................................................... 1
- Prevalence .......................................................................................................................................................... 1
- Etiology, Natural History, and Burden of Disease ............................................................................................... 2
- Risk Factors ....................................................................................................................................................... 3
- Rationale for Screening/Screening Strategies ........................................................................................................ 4
- Interventions/Treatment ...................................................................................................................................... 5
- Current Clinical Practice/Recommendations of Other Groups ............................................................................ 5

## Chapter 2. Methods
- Key Questions and Analytic Framework ........................................................................................................... 6
  - Key Questions .................................................................................................................................................. 6
  - Contextual Questions ....................................................................................................................................... 6
- Search Strategies .................................................................................................................................................. 6
- Study Selection .................................................................................................................................................... 7
- Data Abstraction and Quality Rating .................................................................................................................. 7
- Data Synthesis .................................................................................................................................................... 7
- External Review .................................................................................................................................................. 7
- Response to Public Comments .......................................................................................................................... 8

## Chapter 3. Results
- Key Question 1. What Is the Effectiveness of Screening Versus Not Screening for Celiac Disease in Asymptomatic Adults, Adolescents, or Children on Morbidity, Mortality, or Quality of Life? .............................................................................................................................. 9
- Key Question 2. What Is the Effectiveness of Targeted Versus Universal Screening for Celiac Disease in Asymptomatic Adults, Adolescents, or Children on Morbidity, Mortality, or Quality of Life? ................................................................................................. 9
- Key Question 3. What Are the Harms of Screening for Celiac Disease? .......................................................... 9
- Key Question 4. What Is the Accuracy of Screening Tests for Celiac Disease? .................................................. 9
  - Summary ......................................................................................................................................................... 9
  - Evidence ......................................................................................................................................................... 10
- Key Question 5. Does Treatment of Screen-Detected Celiac Disease Lead to Improved Morbidity, Mortality, or Quality of Life Compared With No Treatment? ............................................................ 11
  - Summary ....................................................................................................................................................... 11
  - Evidence ......................................................................................................................................................... 11
- Key Question 6. Does Treatment of Screen-Detected Celiac Disease Lead to Improved Morbidity, Mortality, or Quality of Life Compared With Treatment Initiated After Clinical Diagnosis? ................................................................. 12
- Key Question 7. What Are the Harms Associated With Treatment of Celiac Disease? ............................... 12
- Contextual Question 2. What Is the Natural History of Subclinical or Silent Celiac Disease? ....................... 14
Chapter 4. Discussion

Summary of Review Findings .................................................................16
Limitations ............................................................................................16
Emerging Issues/Next Steps .................................................................17
Relevance for Priority Populations .......................................................17
Future Research ....................................................................................17
Conclusions ..........................................................................................18

References ............................................................................................19

Figure
Figure. Analytic Framework

Tables
Table 1. Recommendations of Other Groups
Table 2. Natural History of Celiac Disease
Table 3. Summary of Evidence

Appendixes
Appendix A. Detailed Methods
  Appendix A1. Search Strategies
  Appendix A2. Inclusion and Exclusion Criteria
  Appendix A3. Literature Flow Diagram
  Appendix A4. Excluded Studies List
  Appendix A5. U.S. Preventive Services Quality Criteria for Rating Individual Studies

Appendix B. Evidence and Quality Tables
  Appendix B1. Systematic Review of Diagnostic Accuracy Studies
  Appendix B2. Quality Assessment of Systematic Review of Diagnostic Accuracy Studies
  Appendix B3. Diagnostic Accuracy Studies in Asymptomatic Populations
  Appendix B4. Quality Assessment of Diagnostic Accuracy Studies in Asymptomatic Populations
  Appendix B5. Randomized, Controlled Trial of Treatment
  Appendix B6. Quality Assessment of Randomized, Controlled Trial of Treatment
Chapter 1. Introduction

Purpose and Previous U.S. Preventive Services Task Force Recommendation

This report, commissioned by the Agency for Healthcare Research and Quality (AHRQ), will be used by the U.S. Preventive Services Task Force (USPSTF) to develop a recommendation on screening for celiac disease in adults, adolescents, and children age 3 years or older. This topic has not previously been reviewed by the USPSTF.

Condition Definition

Celiac disease is a multisystem autoimmune disorder in genetically predisposed persons that is triggered by ingestion of dietary gluten. Gluten is a protein complex found in wheat, rye, and barley. In persons with celiac disease, ingestion of gluten causes immune-mediated inflammatory damage to the mucosa of the small intestine and subsequent malabsorption of nutrients. Celiac disease can manifest as both gastrointestinal and nongastrointestinal illness. Other names for the disorder include celiac sprue, gluten-sensitive enteropathy, and nontropical sprue.

Prevalence

A challenge in estimating prevalence of celiac disease is that in a number of studies, diagnosis was based on serologic testing without histologic confirmation, potentially overestimating prevalence of celiac disease due to false-positive serologic tests. However, a systematic review of 38 studies in North America and Western Europe found that celiac disease prevalence was 0.15 to 1.87 percent in studies that included biopsy confirmation of positive serologic tests, and was similar (0.15% to 2.67%) in studies that did not perform biopsy confirmation in all patients; among the three U.S. studies, prevalence ranged from 0.40 to 0.95 percent in adults. In the largest multicenter U.S. study included in the systematic review, overall prevalence of celiac disease diagnosed by endomysial antibody (EMA)-positive serology and confirmed by biopsy (<30%) or human leukocyte antigen (HLA) haplotypes DQ2 and DQ8 among 4,126 average-risk persons was 0.75 percent, with prevalence of 0.95 percent among adults, 0.31 percent among children, 0.72 percent among women, and 0.78 percent among men. Prevalence among minority groups was 0.42 percent; results were not presented for specific minority groups. A screening study for celiac disease using stored sera from a population-based sample of adults age 50 years or older in Minnesota found that the prevalence of undiagnosed celiac disease was 0.8 percent, as defined by initial tissue transglutimase (tTG) immunoglobulin (Ig)A testing followed by EMA tests. Median age of those diagnosed was 63 years and 51 percent were women. In a study of 7,798 persons age 6 years or older who participated in the 2009–2010 National Health and Nutrition Examination Survey (NHANES), the prevalence of celiac disease, as defined by positive serology or patient-self report, was 0.71 percent among the general population, 0.76 percent among those age 20 years or older, 0.62 percent among women, and 1.01 percent among
non-Hispanic whites. Some data suggest that the prevalence of celiac disease in the United States has increased over the past several decades for reasons that are not well understood but may be related to changes in dietary gluten. (See Contextual Question 1 for prevalence of celiac disease among patients without overt symptoms.)

Etiology, Natural History, and Burden of Disease

Celiac disease is caused by an immune response to dietary gluten in genetically susceptible persons. Specifically, persons with alleles that encode for HLA-DQ2 and DQ8 proteins are at risk for celiac disease. However, many persons with these alleles do not develop celiac disease, meaning that their presence is necessary but not sufficient for disease. Gliadin, the alcohol-soluble fraction of gluten, triggers both adaptive and innate immune system responses causing infiltration of inflammatory cells into the lamina propria and epithelium of the small intestine, resulting in villous atrophy. Inflammatory injury to the small intestine results in loss of absorptive surface area, reduction in digestive enzymes, and impaired absorption of micronutrients, including fat-soluble vitamins and iron. Although some research suggests an association between breastfeeding with delayed introduction of gluten into the infant diet and decreased risk of celiac disease, more recent literature has not found an association between breastfeeding and risk of celiac disease. Gastrointestinal illness may increase the risk of celiac disease in infancy.

Celiac disease affects both children and adults. Seroconversion to antibodies associated with celiac disease may occur at any time, and disease progression can take place over months or years. Data suggest that the average age at celiac disease diagnosis has increased and is now in the fourth to sixth decades of life.

The clinical presentation, severity of symptoms, and natural history of celiac disease varies among both adults and children. Classic celiac disease presents with symptoms of malabsorption, such as diarrhea, abdominal pain, and weight loss. In children, classic celiac disease is characterized by onset of gastrointestinal symptoms and impaired growth between ages 6 and 24 months, but this is now an uncommon presentation. Analysis of trends among 590 patients with biopsy-diagnosed celiac disease in New York from 1981 to 2004 found that the percentage presenting with diarrhea decreased from 91.3 percent before 1980 to 37.2 percent after 2000, perhaps due to increased awareness of celiac disease, increased screening in asymptomatic or mildly symptomatic persons, and/or ease of serologic testing. Celiac disease now presents more typically with nongastrointestinal, nonspecific manifestations of disease such as anemia, osteoporosis, chronic fatigue, peripheral neuropathy or ataxia, aphthous stomatitis, dermatitis herpetiformis, infertility, recurrent fetal loss, or short stature. Children may also experience pubertal delay and dental enamel defects.

Another form of celiac disease is subclinical disease, or disease that is below the threshold of clinical detection (i.e., without signs or common symptoms sufficient to trigger testing for celiac disease). Persons with subclinical celiac disease may have nonspecific symptoms of celiac disease, such as fatigue, that are not recognized until initiation of a gluten-free diet. Asymptomatic, or silent, celiac disease refers to persons who have been diagnosed with celiac
disease by serologic testing and intestinal biopsy but do not manifest any common symptoms or signs of celiac disease. Potential celiac disease refers to persons with and without symptoms who have positive serology but absent or mild intestinal damage on biopsy. Latent celiac disease, a less commonly used term, describes persons previously diagnosed with celiac disease who have normal intestinal mucosa while on a gluten-free diet or those with normal intestinal mucosa while on a gluten-containing diet who later develop celiac disease. The natural history of subclinical, asymptomatic, potential, and latent celiac disease is not well-defined, and it is not entirely clear if they represent progressive stages of celiac disease or distinct subtypes. In an Italian retrospective study of 549 patients with celiac disease, 45.7 percent showed classical, 47.7 percent subclinical, and 6.6 percent silent forms of celiac disease at the time of diagnosis. (See Contextual Question 2 for additional details regarding the natural history of subclinical or silent celiac disease.)

Some evidence suggests that celiac disease is associated with excess mortality, which is primarily attributed to increased risk for intestinal adenocarcinoma and enteropathy-associated T-cell lymphoma. A recent meta-analysis of observational studies from the United States and Europe showed an increased risk for all-cause mortality in persons with celiac disease (odds ratio [OR], 1.24 [95% confidence interval (CI), 1.19 to 1.30]). In a subgroup analysis, patients identified by positive serology alone were also at an increased risk of all-cause mortality (OR, 1.16 [95% CI, 1.02 to 1.31]) and non-Hodgkin lymphoma (OR, 2.55 [95% CI, 1.02 to 6.36]). However, some data suggest that asymptomatic or silent celiac disease is not associated with increased mortality or other complications of celiac disease. A retrospective study of 549 patients with celiac disease diagnosed by intestinal biopsy found that the rate of complications on a gluten-free diet for a mean duration of 7 years, including malignancy, was highest among those with classic celiac disease (5.58%); no patients with silent disease experienced complications.

Nonceliac gluten sensitivity (NCGS) refers to a condition in which persons with symptoms such as abdominal pain and bloating improve with removal of exposure to gluten but do not have diagnostic features of celiac disease and are not thought to be at increased risk of nutritional deficiency states or other complications associated with celiac disease. Because NCGS is defined based on the presence of symptoms rather than on diagnostic tests, it does not meet criteria for screening and is therefore outside the scope of this review. NCGS is associated with a broad range of symptoms that may manifest as heterogeneous subtypes. A recent double-blinded trial of persons thought to have NCGS found no difference in symptoms following randomization and exposure to high-gluten, low-gluten, or nongluten diets, potentially calling into question the underlying concept for this condition.

**Risk Factors**

A positive family history is a risk factor for celiac disease. The frequency of celiac disease is higher among first- and second-degree relatives of persons with celiac disease, although prevalence estimates range from 5 to 20 percent. Frequency of celiac disease is also higher among persons with other autoimmune diseases, such as type 1 diabetes mellitus, inflammatory luminal gastrointestinal disorders, Down syndrome, Turner’s syndrome, IgA deficiency, and IgA nephropathy.
As discussed previously, celiac disease is more commonly diagnosed among adults ages 40 to 60 years and among non-Hispanic whites. Data regarding risk of celiac disease among women is mixed, but several large-scale prevalence studies found that rates of celiac disease are similar among men and women. The major genetic risk factor for celiac disease is inheritance of HLA-DQ2 and DQ8 alleles, which is more likely among first- and second-degree relatives of persons with diagnosed celiac disease.

**Rationale for Screening/Screening Strategies**

Studies in the United States and Europe suggest that celiac disease may be underdiagnosed, based on the prevalence of positive serologic tests (initial tTG antibody tests followed by EMA testing for those with positive or borderline findings) in persons not previously diagnosed with celiac disease. Evidence also suggests that diagnosis of celiac disease is often delayed. A survey of 1,612 patients with celiac disease in the United States found that symptoms were present for a mean of 11 years before diagnosis. Screening might enable earlier initiation of treatment and reduce the burden of morbidity and mortality associated with untreated celiac disease.

Clinical practice guidelines recommend an algorithmic approach to diagnostic testing for celiac disease, starting with the tTG IgA test and further testing based on the probability of disease. The tTG IgA test is the standard method of testing for celiac disease in persons older than age 2 years. The reported sensitivity of the tTG IgA test is about 95 percent and specificity is 95 percent or greater. In patients in whom celiac disease is suspected but IgA deficiency is a consideration, total IgA levels are measured. Alternatively, IgA testing as well as tTG immunoglobulin G (IgG) and/or IgG deamidated gliadin peptide (DGP) tests can be obtained in such patients. Clinical practice guidelines in the United States and Europe recommend intestinal biopsy to confirm the diagnosis of celiac disease (e.g., based on presence of villous atrophy classified as grade 3 or higher on Marsh criteria) and to distinguish celiac disease from other disorders affecting the small intestine. Intestinal biopsy may also be performed if clinical suspicion for celiac disease is high but serologic tests are negative. It has been suggested that a combination of serologic tests could be used to establish celiac disease diagnosis as an alternative to biopsy, but it is unclear how frequently celiac disease is diagnosed in the absence of biopsy in current clinical practice. Rarely, capsule endoscopy is used to establish a diagnosis of celiac disease in patients who are unwilling or unable to undergo upper endoscopy with intestinal biopsy. HLA-DQ2/DQ8 genotyping is not used routinely to diagnose celiac disease but may be used to rule out the disease in cases with equivocal serologic tests and/or small-bowel histologic findings.

Many persons initiate a gluten-free diet prior to consultation with a health care provider, which complicates the diagnosis of celiac disease and may result in false-negative antibody tests or biopsies. Serologic testing may still be obtained depending on the duration of gluten-free diet, or deferred until gluten has been reintroduced into the diet. HLA-DQ2/DQ8 genotyping is sometimes used to exclude celiac disease before having patients undergo a gluten challenge.

Antigliadin antibodies (AGAs) were previously routinely used to diagnose celiac disease but are
Interventions/Treatment

The mainstay of treatment for celiac disease is lifelong adherence to a gluten-free diet. Short-term vitamin and mineral repletion may also be recommended. Removal of gluten from the diet reverses disease manifestations in a majority of patients. However, complete removal of gluten from the diet is a challenge, as gluten is present in a wide variety of foods, and gluten-free foods can be difficult to obtain and expensive. Nonadherence among patients is also common. A systematic review reported rates of strict adherence to a gluten-free diet of 42 to 91 percent, depending on the definition of adherence and method of ascertainment. Adherence was lowest among ethnic minorities and persons diagnosed in childhood, and rates of adherence were similar among screen-detected and symptomatic patients. Patients who do not respond to a gluten-free diet are often evaluated for concurrent lactose or other carbohydrate intolerance, pancreatic insufficiency, inflammatory bowel disease, and functional gastrointestinal disorders.

Refractory celiac disease occurs in a minority of patients and is characterized by ongoing symptoms of malabsorption despite adherence to a gluten-free diet for 6 to 12 months. These patients may receive treatment with corticosteroids and other immunosuppressive agents such as azathioprine, 6-mercaptopurine, or cyclosporine. Data regarding the effectiveness of these agents is limited to observational studies.

Current Clinical Practice/Recommendations of Other Groups

Clinical practice guidelines recommend testing for celiac disease among persons with signs and symptoms of malabsorption as well as certain populations of asymptomatic persons at increased risk for celiac disease (Table 1). Reliable data on the frequency of screening for celiac disease in clinical practice is not available.

The complex clinical spectrum of celiac disease complicates diagnosis and management. Due to recent media attention to gluten and its potential adverse effects on health, many persons start a gluten-free diet without medical advice. Some experience improvement in gastrointestinal symptoms that are attributed to celiac disease. As discussed previously, clinical improvement on a gluten-free diet is not diagnostic of celiac disease, as many other forms of gluten reaction have been described. Symptomatic improvement may also be due to a placebo effect or to other healthful changes that occur in conjunction with a modified diet.
Chapter 2. Methods

Key Questions and Analytic Framework

Using the methods developed by the USPSTF, the USPSTF and AHRQ determined the scope and key questions for this review. In conjunction with USPSTF members and the AHRQ Medical Officer, investigators created an analytic framework with the key questions and the patient populations, interventions, and outcomes reviewed (Figure).

Key Questions

1. What is the effectiveness of screening versus not screening for celiac disease in asymptomatic adults, adolescents, or children on morbidity, mortality, or quality of life?
2. What is the effectiveness of targeted versus universal screening for celiac disease in asymptomatic adults, adolescents, or children on morbidity, mortality, or quality of life? (Targeted screening refers to testing in patients with family history or other risk factors for celiac disease.)
3. What are the harms of screening for celiac disease?
4. What is the accuracy of screening tests for celiac disease?
5. Does treatment of screen-detected celiac disease lead to improved morbidity, mortality, or quality of life compared with no treatment?
6. Does treatment of screen-detected celiac disease lead to improved morbidity, mortality, or quality of life compared with treatment initiated after clinical diagnosis?
7. What are the harms associated with treatment of celiac disease?

We also addressed two contextual questions requested by the USPSTF to help inform the report. Contextual questions address background areas identified by the USPSTF for informing its recommendations and are not reviewed using systematic review methodology, but rather summarize important contextual evidence.

Contextual Questions

1. Among patients without overt symptoms, what is the prevalence of celiac disease in children, adolescents, and adults in the United States?
2. What is the natural history of subclinical or silent celiac disease?

Search Strategies

We searched the Cochrane Central Register of Controlled Trials and Cochrane Database of Systematic Reviews and Ovid MEDLINE (to mid-June 2016) for relevant studies and systematic reviews. Search strategies are available in Appendix A1. We also reviewed reference lists of relevant articles.
Study Selection

At least two reviewers independently evaluated each study to determine inclusion eligibility. We selected studies on the basis of inclusion and exclusion criteria developed for each key question (Appendix A2). For screening and diagnosis, the population of interest was asymptomatic adults, adolescents, or children age 3 years or older without known celiac disease who had not sought evaluation for potential celiac disease, including persons at higher risk due to family history or presence of conditions associated with celiac disease. For treatment, the population of interest was persons with screen-detected celiac disease who were asymptomatic. We included studies of mildly symptomatic patients if no studies were available in asymptomatic populations. Screening tests were serologic tests or questionnaires. We included randomized trials, cohort studies, and case-control studies performed in primary care or primary care–applicable settings of screening versus no screening, targeted versus universal screening, treatment versus no treatment, and immediate versus delayed treatment that reported morbidity (including outcomes related to nutritional deficiencies, gastrointestinal symptoms), cancer incidence, mood and anxiety, child growth outcomes, infection rates, quality of life, or mortality. For diagnostic accuracy, we included cohort and cross-sectional studies that compared screening tests against endoscopy with biopsy as the reference standard. We excluded studies that focused on intermediate outcomes such as laboratory values for nutritional or other deficiencies and studies that evaluated diagnostic accuracy using a case-control design. To summarize the diagnostic accuracy of screening tests in populations that were not asymptomatic, we included good-quality systematic reviews. The selection of literature is summarized in the literature flow diagram (Appendix A3). Appendix A4 lists excluded studies with reasons for exclusion.

Data Abstraction and Quality Rating

One investigator abstracted details about each article’s study design, patient population, setting, screening method, treatment regimen, analysis, followup, and results. A second investigator reviewed data abstraction for accuracy. Two investigators independently applied criteria developed by the USPSTF to rate the quality of each study as good, fair, or poor (Appendix A5). Discrepancies were resolved through consensus.

Data Synthesis

We assessed the aggregate internal validity (quality) of the body of evidence for each key question (“good,” “fair,” or “poor”) using methods developed by the USPSTF, based on the number, quality and size of studies, consistency of results between studies, and directness of evidence. There were too few studies to perform meta-analysis.

External Review

The draft report was reviewed by content experts, USPSTF members, AHRQ Project Officers, and collaborative partners, and posted for public comment.
Response to Public Comments

The draft report was posted for public comment on the USPSTF Web site from May 3, 2016 to May 30, 2016, and few comments were received. No comments identified missing studies or errors in the evidence reviewed, resulting in no changes to the findings or conclusions of the report.
Chapter 3. Results

Key Question 1. What Is the Effectiveness of Screening Versus Not Screening for Celiac Disease in Asymptomatic Adults, Adolescents, or Children on Morbidity, Mortality, or Quality of Life?

We identified no studies on the effectiveness of screening versus no screening for celiac disease in asymptomatic adults, adolescents, or children on morbidity, mortality, or quality of life.

Key Question 2. What Is the Effectiveness of Targeted Versus Universal Screening for Celiac Disease in Asymptomatic Adults, Adolescents, or Children on Morbidity, Mortality, or Quality of Life?

We identified no studies on the effectiveness of targeted screening of persons with a family history or other risk factors for celiac disease versus universal screening for celiac disease in asymptomatic adults, adolescents, or children on morbidity, mortality, or quality of life.

Key Question 3. What Are the Harms of Screening for Celiac Disease?

We identified no trials on the harms of screening versus no screening for celiac disease.

Key Question 4. What Is the Accuracy of Screening Tests for Celiac Disease?

Summary

One good-quality systematic review found that tTG antibody tests were associated with high sensitivity and specificity in populations not restricted to asymptomatic persons. Based on new studies, the pooled sensitivity in the systematic review was 92.8 percent (95% CI, 90.3% to 94.8%) and specificity was 97.9 percent (95% CI, 96.4% to 98.8%), for a positive likelihood ratio (PLR) of 45.1 (95% CI, 25.1 to 75.5) and negative likelihood ratio (NLR) of 0.07 (95% CI, 0.05 to 0.10). EMA tests were also associated with strong likelihood ratios. Limited evidence from two studies of serologic testing in asymptomatic, high-risk children and younger adults reported lower sensitivity (57% to 71%); specificity ranged from 83 to 98 percent.
Evidence

A recent good-quality systematic review on the diagnostic accuracy of tests for celiac disease included 56 original studies and 12 prior systematic reviews (Appendices B1 and B2). Sample sizes ranged from 62 to more than 12,000 subjects. Three primary studies focused on diagnostic accuracy of testing in children and/or adolescents, six evaluated a mixed population of children and adults, and the remainder focused on adults. One study was conducted in the United States, five studies in the Middle East, one in India, one in Argentina, and the rest in Europe. Tests evaluated included tTG, EMA, DGP, and video capsule endoscopy. Only two studies reported diagnostic accuracy in asymptomatic persons (Appendices B3 and B4).

Overall, including studies of persons with symptoms or in whom symptom status was not described, the systematic review found high strength of evidence that the tTG IgA test was associated with high (>90%) sensitivity and specificity and the EMA IgA test was associated with high specificity, based on consistent results from prior systematic reviews and new studies. For the tTG IgA test, the pooled sensitivity, based on new studies, was 92.8 percent (95% CI, 90.3% to 94.8%) and specificity was 97.9% (95% CI, 96.4% to 98.8%), for a PLR of 45.1 (95% CI, 25.1 to 75.5) and NLR of 0.07 (95% CI, 0.05 to 0.10). For the EMA IgA test, the pooled sensitivity, based on new studies, was 73.0 percent (95% CI, 61.0% to 83.0%) and specificity was 99.0 percent (95% CI, 98.0% to 99.0%), for a PLR of 65.6 (95% CI, 35.6 to 120.8) and NLR of 0.28 (95% CI, 0.17 to 0.41). Results for the DGP IgA test indicated somewhat weaker likelihood ratios; the pooled sensitivity was 87.8 percent (95% CI, 85.6% to 89.9%) and specificity was 94.1 percent (95% CI, 92.5% to 95.5%), for a PLR of 13.3 (95% CI, 9.6 to 18.4) and NLR of 0.12 (95% CI, 0.08 to 0.18). For video capsule endoscopy, the pooled sensitivity was 89.0 percent (95% CI, 82.0% to 94.0%) and specificity was 95.0 percent (95% CI, 89.0% to 99.0%), for a PLR of 12.9 (95% CI, 2.9 to 57.6) and NLR of 0.16 (95% CI, 0.10 to 0.25).

Three studies in the systematic review compared the accuracy of tests by age group. Sensitivity and specificity were generally similar across age groups, with the exception of one study that reported specificity of 26 percent for the DGP IgA test among those age 18 years or younger. Sensitivity was somewhat lower in adults than in children, but differences were slight.

Only two studies included in the systematic review reported diagnostic accuracy in asymptomatic persons (Appendices B3 and B4). A small (n=62), fair-quality study of patients in Iraq with type 1 diabetes mellitus (mean age, 23 years) without symptoms or a family history of celiac disease evaluated tTG IgA, tTG IgG, EMA IgA, AGA IgA, and AGA IgG assays. The prevalence of celiac disease, based on biopsy, was 11.3 percent (7/62); sensitivity ranged from 57 percent for the tTG IgG test to 71 percent for the tTG IgA and EMA IgA tests, resulting in positive predictive values of 50.0 to 71.4 percent; specificity was similar across tests, ranging from 93 to 98 percent, for negative predictive values of 94.4 to 96.4 percent.

Another fair-quality study reported diagnostic accuracy of the combination of tTG IgA and EMA IgA tests in a subgroup of 158 asymptomatic Czech children and adolescents ages 16 months to 19 years at higher risk for celiac disease because they had a first-degree relative with celiac disease or had an associated disease, such as type 1 diabetes mellitus or autoimmune...
thyroiditis. 37 The prevalence of Marsh 2 or 3 small-bowel mucosal villous atrophy was 78.5 percent (124/158), with sensitivity of 67 percent and specificity of 83 percent for the combination of tTG IgA levels more than 10 times the upper limit of normal and a positive EMA IgA test. Results were not reported for the subgroup of patients with Marsh 3 biopsy findings. Sensitivity was 70 percent and specificity was 81 percent for patients with a first-degree relative with celiac disease (n=32); sensitivity was 64 percent and specificity was 93 percent for patients with type 1 diabetes mellitus (n=40).

Key Question 5. Does Treatment of Screen-Detected Celiac Disease Lead to Improved Morbidity, Mortality, or Quality of Life Compared With No Treatment?

Summary

One small (n=40), fair-quality trial of screen-detected, asymptomatic adults found that a gluten-free diet was associated with small improvements in gastrointestinal symptoms (<1 point on a 7-point scale) versus no gluten-free diet after 1 year, but there were no changes on most quality of life outcomes. No other study evaluated the effects of a gluten-free versus no gluten-free diet on clinical outcomes.

Evidence

One fair-quality trial (n=40) evaluated a gluten-free versus normal gluten–containing diet among screen-detected adults who were asymptomatic relatives of persons with celiac disease (Appendices B5 and B6). 59 Median age of participants was 42 years. Diagnosis of celiac disease was based on a positive serum EMA test. Although biopsy was performed, histopathologic findings of celiac disease were not required for study entry, and biopsy results were blinded from study researchers until completion of the trial. At baseline, the mean villous height to crypt depth ratio was 1.0 in the gluten-free diet group and 0.8 in the nongluten-free diet group; two patients in each group had a normal villous height to crypt depth (>2.0).

At 1 year, subjects on a gluten-free diet reported significant improvements in total gastrointestinal symptoms compared with those on a nongluten-free diet, based on the overall Gastrointestinal Symptoms Ratings Scale (difference in mean change, -0.4 on a 7-point scale [95% CI, -0.7 to -0.1]), as well as on the diarrhea (difference in mean change, -0.6 [95% CI, -1.1 to 0.0]), indigestion (difference in mean change, -0.7 [95% CI, -1.1 to -0.2]), and reflux subscales (difference in mean change, -0.5 [95% CI, -0.9 to -0.1]), with no differences on the constipation or abdominal pain subscales. The gluten-free diet group also reported greater improvement on the anxiety subscale of the Psychological General Well-Being Scale (difference in mean change, 1.6 on a 6-point scale [95% CI, 0.4 to 2.8]), with no differences on the depression, well-being, self-control, general health, or vitality subscales. There were no differences in any subscales of the Short-Form 36 Survey aside from social functioning, which was worse in the gluten-free diet group (difference in mean change, -8.3 [95% CI, -15.8 to -0.8]). There were no differences between groups in intermediate outcomes such as mean blood
hemoglobin level, mean serum total iron level, mean body mass index, mean percent total body fat, or mean lumbar spine or femoral neck bone mineral density. After 2 years, more than 90 percent of subjects reported adherence to the gluten-free diet, and improvements in histopathologic findings were observed in the gluten-free diet group at 1 year compared with the nongluten-free diet group.

An earlier, small (n=23) trial conducted at the same center did not meet inclusion criteria. Although it randomized patients identified through EMA testing to a gluten-free or normal diet, 87 percent (20/23) of patients had moderate or severe symptoms. All patients had nondiagnostic (Marsh 1 or 2) histologic findings on small bowel biopsy. Over the course of 1 year, a gluten-free diet was associated with significantly improved subjective clinical symptom ratings, with all patients’ ratings changing from severe/moderate to slight/no symptoms (p<0.05), compared with no changes on a nongluten-free diet.

Three small (n=14 to 32) studies evaluated effects of a gluten-free diet in asymptomatic adults with celiac disease but did not meet inclusion criteria because they did not have a nongluten-free diet control group. Each study evaluated effects before initiation of a gluten-free diet and at 1 to 2 years. Following initiation of a gluten-free diet, one study found worse perceived health and more concern about health, one study found no differences in measures of quality or life or general health, and one study found small improvements in gastrointestinal symptoms but no differences in quality of life.

**Key Question 6. Does Treatment of Screen-Detected Celiac Disease Lead to Improved Morbidity, Mortality, or Quality of Life Compared With Treatment Initiated After Clinical Diagnosis?**

We identified no studies on the effectiveness of treatment of screen-detected celiac disease compared with treatment initiated after clinical diagnosis on morbidity, mortality or quality of life.

**Key Question 7. What Are the Harms Associated With Treatment of Celiac Disease?**

The trial of a gluten-free diet by Kurppa and colleagues (included for key question 5) reported no withdrawals “as a result of major symptoms or complications.” We identified no other study on harms of gluten-free versus nongluten-free diet in persons with screen-detected celiac disease.
Contextual Question 1. Among Patients Without Overt Symptoms, What Is the Prevalence of Celiac Disease in Children, Adolescents, and Adults in the United States?

Reliable data regarding the prevalence of subclinical and silent celiac disease in the United States are not available. Most prevalence studies of the general population were not designed to determine whether participants had symptoms potentially attributable to celiac disease or whether they were truly asymptomatic. In a large (n=7,798) NHANES study of persons age 6 years or older, the prevalence of celiac disease, as defined by a positive tTG IgA and EMA IgA test, was 0.71 percent among the general population, 0.76 percent among those age 20 years or older, 0.62 percent among women, and 1.01 percent among non-Hispanic whites. Study participants were asked whether they had previously been diagnosed with celiac disease and whether they were on a gluten-free diet but were not interviewed regarding symptoms that could be attributed to celiac disease. Other studies of the general adult U.S. population found a celiac disease prevalence of 0.2 to 0.9 percent, based on positive serologic tests, specifically initial tTG antibody tests followed by EMA testing. None of these studies reported whether participants had symptoms that could be caused by celiac disease. Studies from Europe reported the proportion of patients with celiac disease who were asymptomatic. In an Italian retrospective study of 549 patients with celiac disease diagnosed by intestinal biopsy, 45.7 percent presented with classical celiac disease and 6.6 percent were asymptomatic. Another Italian study found that of 770 patients with celiac disease, 79 percent presented with classical celiac disease and 21 percent presented with atypical or silent celiac disease.

Presumably, many cases of celiac disease detected by screening would be subclinical or silent. However, a limitation of many studies is that diagnosis of celiac disease was based on positive results on combinations of serologic tests without histologic confirmation. However, serologic tests are associated with a small proportion of false positives in symptomatic persons. At a given diagnostic accuracy, the positive predictive value of serologic tests will be lower in populations with a lower prevalence of celiac disease.

Even when intestinal biopsy is performed, distinguishing between persons with false-positive serologic findings and persons with subclinical celiac disease can be a challenge, because biopsy findings may be subtle or absent due to patchy disease or inadequate sampling. Most studies reported high concordance between positive serology and intestinal biopsy. However, in a study of 1,461 Estonians ages 15 to 95 years who were screened for celiac disease with AGA IgA testing, 3.5 percent (52 persons) had positive serology, but none were symptomatic or had biopsy results consistent with celiac disease. Among 20 screen-detected adults in Northern Ireland with positive celiac disease serology who agreed to have intestinal biopsy, only three had villous atrophy. Of these, one was asymptomatic and two later endorsed symptoms attributed to celiac disease.
Data regarding the proportion of persons with silent or subclinical celiac disease who later develop symptomatic celiac disease are limited. In a study of stored sera from young adults at Warren Air Force Base collected from 1948 to 1954, none of the 14 subjects with undiagnosed celiac disease, based on serologic tests, received a clinical diagnosis of celiac disease within 45 years of followup. A study of adults in Maryland based on 3,511 matched samples of stored sera from 1974 and 1989 found that among 18 cases diagnosed with celiac disease, based on positive EMA IgA and positive/borderline results for tTG IgA, two persons received a clinical diagnosis of celiac disease at a mean followup of 31.1 years. In a study of 16,847 adults age 50 years or older in Minnesota, 129 were found to have undiagnosed celiac disease, based on positive tTG IgA and EMA IgA tests. During a median followup of 10.3 years, 20 persons were clinically diagnosed with celiac disease. A study of 3,654 Finnish children without known celiac disease found that 1.5 percent (56 children) had positive tTG IgA and EMA IgA or IgG tests. Over 7 years of followup, 37 (about 1%) were diagnosed with celiac disease on the basis of biopsy, of which 10 remained clinically silent. A Dutch study of children ages 2 to 4 years diagnosed with celiac disease based on EMA antibodies and confirmatory biopsy through a screening program found that five of 12 asymptomatic children who did not initiate a gluten-free diet remained asymptomatic after 10 years of followup. The other seven children switched to a gluten-free diet due to the development of symptoms; symptoms resolved after initiation of the diet. Another study found that among children (mean age, 29 months) with potential celiac disease (serology positive/Marsh 0–1 histology), 86 percent (18/21) who continued a gluten-containing diet became antibody negative, 9 percent (2/21) had fluctuating antibodies, and 5 percent (1/21) developed overt celiac disease.

Evidence is conflicting whether persons diagnosed with subclinical or silent celiac disease experience the same mortality risk as the general population. The Warren Air Force Base study discussed above found that all-cause mortality was higher among persons with undiagnosed celiac disease (based on positive serology) after 45 years of followup than among seronegative controls within the same cohort. However, symptom status of persons with undiagnosed celiac disease was not reported. In a study of stored sera from German adults collected from 1989 to 1990, positive celiac disease serology was associated with increased risk of all-cause mortality compared with age- and sex-matched controls. Participants were asked about their general self-rated health status, but as in the other study, the prevalence of symptoms attributable to celiac disease was not reported.

A meta-analysis of observational studies reported somewhat conflicting results regarding effects of celiac disease diagnosed by serologic testing and association with increased risk of all-cause mortality and cancer compared with seronegative age- and sex-matched controls. In three studies, screen-detected celiac disease (diagnosed by serologic tests alone, symptoms not reported) was not associated with increased risk of all-cause or cancer mortality compared with age- and sex-matched controls. However, a fourth study found that latent celiac disease (positive serology and normal mucosa) was associated with an estimated excess mortality of 1.7 deaths per 1,000 person-years compared with age- and sex-matched controls in the general population.
population (hazard ratio, 1.35 [95% CI, 1.14 to 1.58]). Symptom status was not reported, but the authors noted that clinical suspicion for celiac disease was the only major indication for small intestinal biopsy in Sweden, suggesting that persons may have been symptomatic. In another study of screen-detected celiac disease among adults in Northern Ireland, positive serologic tests for celiac disease were not associated with excess mortality risk compared with age-specific mortality in the general population.

Some data suggest that subclinical or silent celiac disease is associated with lower risk of developing celiac disease complications than symptomatic disease (Table 2). An Italian retrospective study of 549 patients with celiac disease diagnosed by intestinal biopsy found that the complication rate among patients on a gluten-free diet (mean duration, 7 years [range, 1 to 15 years]) was 5.58 percent in those with classical celiac disease (n=251) and 1.53 percent in those with subclinical celiac disease (n=262, defined as the presence of gluten-sensitive enteropathy on biopsy with extraintestinal but no gastrointestinal symptoms). Complications included gastrointestinal adenocarcinoma, Sjögren’s syndrome, jejunal enteropathy-associated T-cell lymphoma, myocardial infarction, sclerosing cholangitis, herpetiform dermatitis, gastric mucosa-associated lymphoid tissue lymphoma, ulcerative jejunitis, severe nonalcoholic steatohepatitis, recurrent abortion, and autoimmune thrombocytopenia. There was no statistical difference between the mean age of the two groups developing complications. No patient with silent disease (gluten-sensitive enteropathy on biopsy without symptoms) experienced complications. Another Italian study of 770 patients diagnosed with celiac disease (histologic confirmation) evaluated presentation patterns of patients who developed complicated versus noncomplicated celiac disease (p<0.001). Six patients with classic malabsorption symptoms at presentation developed complications compared with no patients with atypical and subclinical celiac disease over a mean of 5 years (p<0.001). Complications included enteropathy-associated T-cell lymphoma, small bowel carcinoma, and refractory celiac disease.
Chapter 4. Discussion

Summary of Review Findings

Table 3 summarizes the evidence reviewed for this update. We identified no studies of screening versus no screening for celiac disease in the target populations for this review (adults, adolescents, and children age 3 years or older). Although serologic tests for celiac disease used in screening appear to be highly accurate, almost all studies on diagnostic accuracy evaluated populations with symptoms of celiac disease or whose symptom status was not reported. Two studies that specifically evaluated patients at high risk for celiac disease based on family history or presence of conditions associated with celiac disease reported lower sensitivity and inconsistent specificity.37, 40

Only one randomized trial evaluated the effectiveness of gluten-free versus no gluten-free diet in asymptomatic persons with screen-detected celiac disease.59 It found that initiation of a gluten-free diet in screen-detected, asymptomatic adults was associated with improved gastrointestinal symptoms, though effects were relatively small (<1 point on a 7-point scale). There were no effects on most measures of quality of life; no harms resulting in withdrawal from the diet occurred. In this study, patients had a first-degree relative with celiac disease and were diagnosed on the basis of serologic testing. Histologic findings of celiac disease were not required for entry, though most patients had some degree of villous atrophy at baseline. Nonetheless, it is possible that this trial could have underestimated benefits of treatment for histologic-proven celiac disease. Three small studies on effects of a gluten-free diet in persons with asymptomatic celiac disease were excluded because they did not include a gluten-containing diet control group.61-63 There were no clear effects on quality of life, though one study62 found increased worry about health following initiation of a gluten-free diet and one study63 reported small improvements in gastrointestinal symptoms.

No study compared the effectiveness of targeted versus universal screening or evaluated effects of immediate initiation of a gluten-free diet versus delaying until the development of symptoms in asymptomatic persons diagnosed with celiac disease.

Limitations

The major limitation of this review is the lack of evidence to address the key questions. There were no studies on screening versus no screening, only two studies on diagnostic accuracy of serologic testing in asymptomatic populations, and only one trial of treatment in asymptomatic, screen-detected persons with celiac disease. Although numerous studies evaluated the diagnostic accuracy of tests for celiac disease in patients who were not asymptomatic, the applicability of findings to screening settings is uncertain. Meta-analysis was not possible, and we could not formally assess for publication bias. We restricted inclusion to English-language articles but found no non-English-language articles on benefits or harms of screening or treatment that appeared to meet inclusion criteria. Although some non-English-language articles assessed diagnostic accuracy, none were clearly conducted in asymptomatic populations.
Emerging Issues/Next Steps

An emerging issue is the development and uptake of methods for diagnosing celiac disease that do not require histologic confirmation. The proportion of patients who are diagnosed with celiac disease or initiate a gluten-free diet based on serologic testing alone is unknown but may be increasing in clinical practice, despite clinical practice guideline recommendations for histologic confirmation.

A related issue is how to classify and manage persons with positive serologic findings but negative or nondiagnostic findings on biopsy. The likelihood that such patients will go on to develop overt celiac disease requires further investigation, and has important implications for management.

A recent randomized trial that screened persons with a first- or second-degree relative with celiac disease and randomized patients to immediate notification and initiation of a gluten-free diet versus no notification or initiation of a gluten-free diet was terminated. We were unable to determine reasons for study termination.

Although there continues to be research on pharmacologic treatments for celiac disease, such treatments are considered an adjunct to a gluten-free diet, which remains the mainstay of therapy.

Relevance for Priority Populations

In the United States, celiac disease is uncommon among racial/ethnic minorities, although it does occur. In an NHANES study, the prevalence of tTG IgA test results were 0.8 percent (27/3,430) among non-Hispanic whites, 0.07 percent (1/1,394) among non-Hispanic blacks, 0.03 percent (1/2,519) among other Hispanics/not Mexican Americans, and 0.2 percent (1/455) among other race/ethnicity categories.

The only randomized trial of treatment with a gluten-free diet among asymptomatic screen-detected persons was restricted to persons younger than age 18 years or older than age 75 years. Although celiac disease is most commonly diagnosed between ages 40 and 60 years, it can affect adolescents and children as well as older adults.

Future Research

Additional research is needed to address all of the key questions addressed in this report. For screening, trials of screening versus no screening that evaluate clinical outcomes are needed. Trials that target high-risk populations, based on family history or presence of conditions associated with celiac disease, would be likely to provide a higher yield of screen-detected persons than trials that screen lower- or average-risk persons, and might be more informative for an initial screening study. Additional studies are needed to determine the accuracy of serologic testing in asymptomatic persons. Trials are also needed on the effects of initiation of a gluten-free diet versus no gluten-free diet in screen-detected persons, and the effects of immediate
initiation on diagnosis versus delayed initiation until the development of symptoms. The in-progress Celiac Disease and Diabetes-Dietary Intervention and Evaluation Trial (Celiac Disease-DIET), which involves screening for asymptomatic celiac disease in children and adults with type 1 diabetes mellitus followed by randomization to a gluten-free or no gluten-free diet, is designed to assess outcomes (including diabetes control, bone mineral density, and health-related quality of life) over 1 year, and should help clarify effects of screening in higher-risk persons. Ideally, future studies would be carried out long enough to determine effects on long-term outcomes related to nutritional deficiencies such as osteoporotic fractures, cancer, and mortality. Because of the uncertain natural history of positive serologic findings without histologic changes, trials should focus on patients with histologic findings of celiac disease, or report analyses stratified according to baseline histologic findings. Trials should evaluate populations across the age spectrum, including children, adolescents, and adults, as celiac disease can be diagnosed in any of these age groups.

Research is also needed to better understand the natural history of subclinical and silent celiac disease, including the proportion of patients who develop symptoms, the proportion who develop complications, and the proportion in whom serologic and/or histologic findings resolve without treatment.

Conclusions

More research is needed to understand the effectiveness of screening for and treatment of celiac disease in asymptomatic adults, adolescents, and children, and optimal screening strategies.
References


Screening for Celiac Disease 21 Pacific Northwest EPC


79. Makharia G. Effect of addition of short course of prednisolone to gluten free diet in naive celiac disease patients (celiac disease). In: ClinicalTrials.gov [Internet]. Bethesda (MD):


Figure. Analytic Framework

Abbreviation: KQ = key question.
### Table 1. Recommendations of Other Groups

<table>
<thead>
<tr>
<th>Organization</th>
<th>Screening/Testing Recommendation for Celiac Disease</th>
</tr>
</thead>
</table>
| American College of Gastroenterology<sup>26</sup>                           | • Persons with signs/symptoms of malabsorption  
• Symptomatic persons with type 1 diabetes mellitus  
• Asymptomatic persons with elevated serum aminotransferase  
• Symptomatic and asymptomatic first-degree relatives of patients with celiac disease                                                                                     |
| National Institute for Health and Care Excellence, United Kingdom<sup>31</sup> | • Persons with any of the following:  
  o Persistent unexplained abdominal or gastrointestinal symptoms  
  o Faltering growth  
  o Prolonged fatigue  
  o Unexpected weight loss  
  o Severe or persistent mouth ulcers  
  o Unexplained iron, vitamin B12, or folate deficiency  
  o Type 1 diabetes, at diagnosis  
  o Autoimmune thyroid disease, at diagnosis  
  o Irritable bowel syndrome (in adults)  
• First-degree relatives of persons with celiac disease  
• Consider serologic testing for persons with any of the following:  
  o Metabolic bone disorder (reduced bone mineral density or osteomalacia)  
  o Unexplained neurological symptoms (particularly peripheral neuropathy or ataxia)  
  o Unexplained subfertility or recurrent miscarriage  
  o Persistently raised liver enzymes with unknown cause  
  o Dental enamel defects  
  o Down syndrome  
  o Turner syndrome  |
| North American Society for Pediatric Gastroenterology, Hepatology, and Nutrition<sup>33</sup> | • Asymptomatic children age ≥3 years with type 1 diabetes mellitus, autoimmune thyroiditis, Down syndrome, Turner syndrome, Williams syndrome, and selective IgA deficiency  
• Asymptomatic children age ≥3 years who are first-degree relatives of patients with celiac disease  
• Children with failure to thrive, persistent diarrhea, and other gastrointestinal symptoms  
• Children with dermatitis herpetiformis, dental enamel hypoplasia of permanent teeth, osteoporosis, short stature, delayed puberty, and iron-deficiency anemia resistant to oral iron |
| Ontario Health Technology Advisory Committee<sup>32</sup>                     | • Persons with signs/symptoms of malabsorption  
• Persons with unexplained iron-deficiency anemia unresponsive to iron supplementation  
• Persons with dermatitis herpetiformis  |

**Abbreviation:** IgA=immunoglobin A.
<table>
<thead>
<tr>
<th>Author, year</th>
<th>Population Country</th>
<th>N</th>
<th>Duration of followup</th>
<th>Definition of CD</th>
<th>Prevalence</th>
<th>Health outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Systematic review, meta-analysis (17 studies)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tio, 2012</td>
<td>Symptomatic and screen-detected CD patients U.S. and Europe N=313,827 Mean age NR</td>
<td>NR</td>
<td>Varied</td>
<td>NR NR</td>
<td>All-cause mortality: OR, 1.24 (95% CI, 1.19-1.30) Mortality from non-Hodgkin lymphoma: OR, 2.61 (95% CI, 2.04-3.33)</td>
<td>All-cause mortality: OR, 1.16 (95% CI, 1.02-1.31) Mortality from non-Hodgkin lymphoma: OR, 2.55 (95% CI, 1.02-6.36)</td>
</tr>
<tr>
<td>Retrospective cohort studies with comparison groups</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Canavan, 2011</td>
<td>Population-based sample of adults from 1990-1995 United Kingdom N=7,527 Mean age NR, range 45-76 years</td>
<td>Median, 16.8 years</td>
<td>Positive IgA EMA</td>
<td>NA 1.16%</td>
<td>NA All-cause mortality was 9.4 deaths per 1,000 person-years (95% CI, 5.4-16.1) After adjustment for age, sex, smoking, and socioeconomic status: 0.98 (95% CI, 0.57-1.69)</td>
<td></td>
</tr>
<tr>
<td>Godfrey, 2010</td>
<td>Population-based sample of adults from 1995-2001 U.S., Minnesota N=16,886 Median age 63, range 52-88 years</td>
<td>10.3 years</td>
<td>Positive tTG IgA and positive EMA IgA</td>
<td>NA 0.8%</td>
<td>NA Hazard ratio for all-cause mortality: 0.8 (95% CI, 0.45-1.41) Hazard ratio for cancer mortality: 0.63 (95% CI, 0.16-2.48)</td>
<td></td>
</tr>
<tr>
<td>Johnston, 1998</td>
<td>Population-based sample of adults, 1983 Northern Ireland N=1,204 Mean age NR</td>
<td>Mean, 11.6 years (range, 11.3-11.9)</td>
<td>Positive IgA gliadin antibody, IgA antireticulin antibody, or EMA IgA</td>
<td>NA 8.47%</td>
<td>NA Relative risk of all-cause mortality: 0.92 (95% CI, 0.5-1.6) Relative risk of cancer mortality: 0.94 (95% CI, 0.3-2.4)</td>
<td></td>
</tr>
<tr>
<td>Lohi, 2009</td>
<td>Population-based sample of adults, 1978-1980 Finland N=6,987 Mean age 51, range 30-95 years</td>
<td>Up to 28 years</td>
<td>Positive tTG IgA or EMA IgA</td>
<td>NA 1.1% EMA positive, 2.9% tTG positive</td>
<td>NA Age- and sex-adjusted relative risk of overall mortality with positive EMA IgA: 0.78 (95% CI, 0.52-1.18) Age- and sex-adjusted relative risk of overall mortality with positive tTG IgA: 1.19 (95% CI, 0.99-1.42)</td>
<td></td>
</tr>
</tbody>
</table>
Table 2. Natural History of Celiac Disease

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Population Country</th>
<th>N</th>
<th>Age</th>
<th>Duration of followup</th>
<th>Definition of CD</th>
<th>Prevalence</th>
<th>Health outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ludvigsson, 2009</td>
<td>Adults who had small intestinal biopsy with CD or latent CD Sweden N=46,121 Median age 30 (with CD) and 36 (with latent CD)</td>
<td>74</td>
<td></td>
<td>8.8 years (with CD), 6.7 years (with latent CD)</td>
<td>Villous atrophy on small intestinal biopsy</td>
<td>NR</td>
<td>Hazard ratio for all-cause mortality in CD: 1.39 (95% CI, 1.33-1.45) Hazard ratio for all-cause mortality in latent CD: 1.35 (95% CI, 1.14-1.58)</td>
</tr>
<tr>
<td>Metzger, 2006</td>
<td>Population-based sample of adults from 1989-1990 Southern Germany N=4,633 Mean age men, 57 years Mean age women, 53 years</td>
<td>71</td>
<td></td>
<td>Median, 7.95 years (range, 11 days-8.9 years)</td>
<td>Positive tTG IgA test</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Rubio-Tapia, 2009</td>
<td>Healthy adults U.S., Warren AFB N=9,133 Mean age 21 years</td>
<td>5</td>
<td></td>
<td>45 years</td>
<td>Positive tTG IgA or EMA IgA</td>
<td>NA</td>
<td>Hazard ratio for all-cause mortality: 3.9 (95% CI, 2.0-7.5)</td>
</tr>
<tr>
<td>Tursi, 2009</td>
<td>CD patients on gluten-free diet enrolled 1993-2006 Italy N=549 Mean age NR</td>
<td>19</td>
<td></td>
<td>NR</td>
<td>Positive small bowel biopsy</td>
<td>45.7% 47.7% subclinical* 6.6% silent</td>
<td>79% 79% 0% silent</td>
</tr>
<tr>
<td>Volta, 2014</td>
<td>Adults diagnosed with CD 1998-2012 Italy N=770 Median age 36 years</td>
<td>65</td>
<td></td>
<td>Mean, 5 years (range, 18 months-14 years)</td>
<td>Varied (combination of duodenal biopsy, serology, and HLA typing based on patient-specific factors)</td>
<td>79% 21%</td>
<td>Rate of complications (enteropathy-associated T-cell lymphoma, small bowel carcinoma, and refractory CD): 0.9%† Rate of complications (enteropathy-associated T-cell lymphoma, small bowel carcinoma, and refractory CD): 0%†</td>
</tr>
</tbody>
</table>

* Subclinical defined by presence of gluten-sensitive enteropathy with extraintestinal symptoms and no gastrointestinal symptoms.
† Difference between groups: p<0.001.

**Abbreviations:** AFB=Air Force Base; CD=celiac disease; CI=confidence interval; EMA=anti-endomysial antibody; HLA=human leukocyte antigen; IgA=immunoglobin A; NA=not applicable; NR=not reported; OR=odds ratio; tTG=tissue transglutaminase.
# Table 3. Summary of Evidence

<table>
<thead>
<tr>
<th>Included studies</th>
<th>Summary of findings</th>
<th>Consistency</th>
<th>Applicability</th>
<th>Limitations</th>
<th>Overall quality</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Key Question 1. What is the effectiveness of screening versus not screening for celiac disease in asymptomatic adults, adolescents, or children on morbidity, mortality, or quality of life?</strong></td>
<td>No studies</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Key Question 2. What is the effectiveness of targeted versus universal screening for celiac disease in asymptomatic adults, adolescents, or children on morbidity, mortality, or quality of life? (Targeted screening refers to testing in patients with family history or other risk factors for celiac disease.)</strong></td>
<td>No studies</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Key Question 3. What are the harms of screening for celiac disease?</strong></td>
<td>No studies</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Key Question 4. What is the accuracy of screening tests for celiac disease?</strong></td>
<td>1 systematic review (of 56 studies and 12 other systematic reviews) One good-quality systematic review found that tTG antibody tests were associated with high sensitivity and specificity in populations not restricted to asymptomatic persons. Based on new studies, the pooled sensitivity in the systematic review was 92.8% (95% CI, 90.3% to 94.8%) and specificity was 97.9% (95% CI, 96.4% to 98.8%), for a positive likelihood ratio of 45.1 (95% CI, 25.1 to 75.5) and negative likelihood ratio of 0.07 (95% CI, 0.05 to 0.10). EMA antibody tests were also associated with strong likelihood ratios.</td>
<td>Consistent</td>
<td>Moderate</td>
<td>Only 2 studies are of asymptomatic persons</td>
<td>Fair</td>
</tr>
<tr>
<td>2 studies (n=220) conducted in asymptomatic persons</td>
<td>Limited evidence from 2 studies of serologic testing in asymptomatic, high-risk children and younger adults reported lower sensitivity (57% to 71%); specificity ranged from 83% to 98%.</td>
<td>-</td>
<td>High Non-U.S. setting</td>
<td>Imprecision</td>
<td>Poor</td>
</tr>
<tr>
<td><strong>Key Question 5. Does treatment of screen-detected celiac disease lead to improved morbidity, mortality, or quality of life compared with no treatment?</strong></td>
<td>1 trial (n=40 randomized from screening pool of 3,031) One small (n=40), fair-quality trial of screen-detected, asymptomatic adults found that a gluten-free diet was associated with small improvements in gastrointestinal symptoms (&lt;1 point on a 7-point scale) versus no gluten-free diet after 1 year, but there were no changes on most quality of life outcomes.</td>
<td>-</td>
<td>High Non-U.S. setting</td>
<td>Imprecision</td>
<td>Poor</td>
</tr>
<tr>
<td><strong>Key Question 6. Does treatment of screen-detected celiac disease lead to improved morbidity, mortality, or quality of life compared with treatment initiated after clinical diagnosis?</strong></td>
<td>No studies</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Key Question 7. What are the harms associated with treatment of celiac disease?</strong></td>
<td>1 trial (n=40 randomized from screening pool of 3,031) The trial included for key question 5 reported no withdrawals &quot;as a result of major symptoms or complications.&quot; We identified no other study on harms of gluten-free vs. nongluten-free diet in persons with screen-detected celiac disease.</td>
<td>-</td>
<td>High Non-U.S. setting</td>
<td>Imprecision</td>
<td>Poor</td>
</tr>
</tbody>
</table>

**Abbreviations:** CI=confidence interval; EMA=endomysial antibody; tTG=tissue transglutaminase.
Appendix A1. Search Strategies

Screening Effectiveness and Harms

Database: Ovid MEDLINE and Ovid OLDMEDLINE
1 Celiac Disease/
2 (celiac adj1 (disease or sprue)).mp.
3 1 or 2
4 Mass Screening/
5 3 and 4
6 screening.ti,ab.
7 3 and 6
8 5 or 7
9 limit 8 to humans
10 limit 9 to English language
11 limit 9 to abstracts
12 10 or 11
13 limit 12 to (clinical trial, all or comparative study or controlled clinical trial or randomized controlled trial)
14 12 and (random$ or control$ or cohort).mp.
15 13 or 14
16 meta-analysis.mp. or exp Meta-Analysis/
17 (cochrane or medline).tw.
18 search$.tw.
19 16 or 17 or 18
20 "Review Literature as Topic"/ or systematic review.mp.
21 19 or 20
22 12 and 21
23 limit 12 to (meta analysis or systematic reviews)
24 limit 12 to evidence based medicine reviews
25 or/22-24
26 15 or 25

Database: EBM Reviews - Cochrane Central Register of Controlled Trials
1 Celiac Disease/
2 (celiac adj1 (disease or sprue)).mp.
3 1 or 2
4 Mass Screening/
5 3 and 4
6 screening.ti,ab.
7 3 and 6
8 5 or 7
9 limit 8 to English language

Diagnostic Accuracy

Database: Ovid MEDLINE and Ovid OLDMEDLINE
1 Celiac Disease/
2 (celiac adj1 (disease or sprue)).mp.
3 1 or 2
4 Immunoglobulin A/
5 Transglutaminases/
6 (IgA or TTG).mp.
7 or/4-6
8 3 and 7
9 8 and screen$.mp.
10 "Sensitivity and Specificity"/
11 (specificity or accurac$ or "predictive value").tw.
Appendix A1. Search Strategies

12 (sensitiv$ or diagnostic).mp.
13 or/10-12
14 3 and 13
15 14 and screen$.mp.
16 9 or 15
17 limit 16 to English language
18 limit 16 to abstracts
19 17 or 18
20 limit 19 to humans

Database: EBM Reviews - Cochrane Central Register of Controlled Trials

1 Celiac Disease/
2 (celiac adj1 (disease or sprue)).mp.
3 1 or 2
4 Immunoglobulin A/
5 Transglutaminases/
6 (IgA or TTG).mp.
7 or/4-6
8 3 and 7
9 8 and screen$.mp.
10 "Sensitivity and Specificity"/
11 (specificity or accurac$ or "predictive value").tw.
12 (sensitiv$ or diagnostic).mp.
13 or/10-12 (14 3 and 13
15 14 and screen$.mp.
16 9 or 15
17 limit 16 to English language
18 limit 16 to abstracts
19 17 or 18

Treatment Effectiveness and Harms

Database: Ovid MEDLINE and Ovid OLDMEDLINE

1 Celiac Disease/dh, dt, pc, th [Diet Therapy, Drug Therapy, Prevention & Control, Therapy]
2 (celiac adj1 (disease or sprue)).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]
3 2 and (dh or dt or pc or th).fs.
4 1 or 3
5 Diet, Gluten-Free/
6 Celiac Disease/
7 5 and 6
8 4 or 7
9 limit 8 to (clinical trial or comparative study or controlled clinical trial or randomized controlled trial)
10 8 and (random$ or control$ or cohort).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]
11 9 or 10
12 limit 8 to (meta analysis or systematic reviews)
13 limit 8 to evidence based medicine reviews
14 meta-analysis.mp. or exp Meta-Analysis/
15 (cochrane or medline).tw.
16 search$.tw.
17 14 or 15 or 16
Appendix A1. Search Strategies

18 "Review Literature as Topic"/ or systematic review.mp.
19 17 or 18
20 8 and 19
21 11 or 12 or 13 or 20
22 limit 21 to English language
23 limit 21 to abstracts
24 22 or 23
25 limit 24 to humans

Database: EBM Reviews - Cochrane Central Register of Controlled Trials
1 Celiac Disease/
2 (celiac adj1 (disease or sprue)).mp.
3 Diet, Gluten-Free/
4 1 or 2 or 3

Systematic Reviews (all Key Questions)

Databases: EBM Reviews - Cochrane Database of Systematic Reviews, EBM Reviews - ACP Journal Club, EBM Reviews - Database of Abstracts of Reviews of Effects, EBM Reviews - Cochrane Central Register of Controlled Trials, EBM Reviews - Cochrane Methodology Register, EBM Reviews - Health Technology Assessment, EBM Reviews - NHS Economic Evaluation Database
1 (celiac or coeliac).ti.
2 1 and gluten.mp.
### Appendix A2. Inclusion and Exclusion Criteria

<table>
<thead>
<tr>
<th>Include</th>
<th>Exclude</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Populations</strong></td>
<td></td>
</tr>
<tr>
<td>KQs 1–3: Asymptomatic adults, adolescents, or children age ≥3 years</td>
<td>KQs 1–3: Symptomatic persons seeking evaluation for potential celiac</td>
</tr>
<tr>
<td>without known celiac disease who have not sought evaluation for</td>
<td>disease</td>
</tr>
<tr>
<td>potential celiac disease (some “asymptomatic” individuals may have</td>
<td></td>
</tr>
<tr>
<td>mild, nonspecific symptoms); studies of asymptomatic patients at</td>
<td></td>
</tr>
<tr>
<td>higher risk (including patients with type 1 diabetes)</td>
<td></td>
</tr>
<tr>
<td>KQ 4: Asymptomatic adults, adolescents, or children age ≥3 years</td>
<td></td>
</tr>
<tr>
<td>without known celiac disease; studies of asymptomatic patients at</td>
<td></td>
</tr>
<tr>
<td>higher risk (including patients with type 1 diabetes)</td>
<td></td>
</tr>
<tr>
<td>KQs 5–7: Patients with screen-detected celiac disease (if evidence in</td>
<td></td>
</tr>
<tr>
<td>such patients is unavailable or very limited, patients with mild celiac</td>
<td></td>
</tr>
<tr>
<td>disease will be included); studies of asymptomatic patients at higher</td>
<td></td>
</tr>
<tr>
<td>risk (including patients with type 1 diabetes)</td>
<td></td>
</tr>
<tr>
<td><strong>Interventions</strong></td>
<td></td>
</tr>
<tr>
<td>KQs 1, 2: Serologic screening (tTG IgA or other commonly used tests)</td>
<td>KQ 4: Screening with biopsy only in patients with positive serology</td>
</tr>
<tr>
<td>KQ 3: Serologic screening (tTG IgA or other commonly used tests);</td>
<td></td>
</tr>
<tr>
<td>diagnostic testing</td>
<td></td>
</tr>
<tr>
<td>KQ 4: Serologic screening (tTG IgA or other commonly used tests);</td>
<td></td>
</tr>
<tr>
<td>questionnaires</td>
<td></td>
</tr>
<tr>
<td>KQs 5–7: Gluten-free diet</td>
<td></td>
</tr>
<tr>
<td><strong>Comparators</strong></td>
<td></td>
</tr>
<tr>
<td>KQ 1: Screening vs. no screening</td>
<td></td>
</tr>
<tr>
<td>KQ 2: Targeted vs. universal screening</td>
<td></td>
</tr>
<tr>
<td>KQ 4: Endoscopy with biopsy</td>
<td></td>
</tr>
<tr>
<td>KQ 5: Screen-detected treatment vs. no treatment</td>
<td></td>
</tr>
<tr>
<td>KQ 6: Screen-detected celiac disease vs. disease detected after</td>
<td></td>
</tr>
<tr>
<td>clinical diagnosis</td>
<td></td>
</tr>
<tr>
<td><strong>Outcomes</strong></td>
<td></td>
</tr>
<tr>
<td>KQs 1, 2, 5, 6: Morbidity (including outcomes related to nutritional</td>
<td>KQs 1, 2, 5, 6: Laboratory values for nutritional or other deficiencies</td>
</tr>
<tr>
<td>deficiencies, such as symptomatic or severe anemia [i.e., requiring</td>
<td></td>
</tr>
<tr>
<td>treatment]); gastrointestinal outcomes (e.g., diarrhea, cramping,</td>
<td></td>
</tr>
<tr>
<td>bloating), cancer incidence, mood and anxiety disorders, child growth</td>
<td></td>
</tr>
<tr>
<td>outcomes, infection rates, and quality of life; mortality</td>
<td></td>
</tr>
<tr>
<td>KQ 3: Labeling, complications/harms from workup/biopsy, and</td>
<td></td>
</tr>
<tr>
<td>overdagnosis</td>
<td></td>
</tr>
<tr>
<td>KQ 4: Sensitivity, specificity, positive and negative predictive values,</td>
<td></td>
</tr>
<tr>
<td>area under the receiver operating curve, and other measures of</td>
<td></td>
</tr>
<tr>
<td>diagnostic test accuracy</td>
<td></td>
</tr>
<tr>
<td>KQ 7: Any harms of treatment</td>
<td></td>
</tr>
<tr>
<td><strong>Settings</strong></td>
<td></td>
</tr>
<tr>
<td>KQs 1–3: Primary care</td>
<td>KQs 1–3: Specialty clinics</td>
</tr>
<tr>
<td><strong>Study designs</strong></td>
<td></td>
</tr>
<tr>
<td>KQs 1–3, 7: Randomized, controlled trials; controlled observational</td>
<td>KQ 4: Case-control studies</td>
</tr>
<tr>
<td>studies; systematic reviews</td>
<td></td>
</tr>
<tr>
<td>KQ 4: Studies evaluating diagnostic accuracy of serologic screening or</td>
<td></td>
</tr>
<tr>
<td>questionnaires compared with intestinal biopsy; systematic reviews</td>
<td></td>
</tr>
<tr>
<td>KQs 5, 6: Randomized, controlled trials; systematic reviews</td>
<td></td>
</tr>
<tr>
<td><strong>Abbreviations:</strong> IgA=immunoglobulin A; KQ=key question; tTG= tissue</td>
<td></td>
</tr>
<tr>
<td>transglutaminase.</td>
<td></td>
</tr>
</tbody>
</table>
Appendix A3. Literature Flow Diagram

Abstracts of potentially relevant articles identified through MEDLINE and Cochrane* databases and other sources†: 3,036

Excluded abstracts and background articles: 2,819

Full-text articles reviewed for relevance to Key Questions: 217

Articles excluded total: 213
- Wrong population: 32
- Wrong intervention: 7
- Wrong outcome: 2
- Wrong comparison: 23
- Wrong study design for Key Question: 77
- Not a study (letter, editorial, nonsystematic review article): 15
- Systematic review used as source document only to identify individual studies: 2
- Study covered in included systematic review: 55

Included studies‡: 4

KQ 1. Screening Effectiveness: No studies
KQ 3. Screening Harms: No studies
KQ 4. Diagnostic Accuracy: 1 systematic review (including 2 studies of asymptomatic persons)
KQ 5. Treatment Effectiveness: 1 trial
KQ 6. Treatment Timing: No studies
KQ 7. Treatment Harms: 1 trial

*Cochrane databases include the Cochrane Central Register of Controlled Trials and the Cochrane Database of Systematic Reviews.
†Other sources include prior reports, reference lists of relevant articles, and systematic reviews

Screening for Celiac Disease

33 Pacific Northwest EPC
Appendix A4. Excluded Studies List


Appendix A4. Excluded Studies List


Grodzinsky E, Hed J, Skogh T. IgA antiendomysium antibodies have a high positive predictive value for celiac disease in asymptomatic patients. Allergy. 1994;49(8):593-7. Excluded: Wrong study design for Key Question.


Appendix A4. Excluded Studies List


Appendix A4. Excluded Studies List


Lewis NR, Scott BB. Systematic review: the use of serology to exclude or diagnose celiac disease (a comparison of the endomysial and tissue transglutaminase antibody tests). Aliment Pharmacol Ther. 2006;24(1):47-54. Excluded: Systematic review or meta-analysis used as source document only to identify individual studies.


Appendix A4. Excluded Studies List


Medical Advisory S. Clinical utility of serologic testing for celiac disease in asymptomatic patients: an evidence-based analysis (Structured abstract). Health Technology Assessment Database. 2014(2). Excluded: Not a study (letter, editorial, non-systematic review article, no original data).

Medical Advisory S. Clinical utility of serologic testing for celiac disease in Ontario (symptomatic patients) (Structured abstract). Health Technology Assessment Database. 2014(2). Excluded: Not a study (letter, editorial, non-systematic review article, no original data).


Appendix A4. Excluded Studies List


Appendix A4. Excluded Studies List


Appendix A4. Excluded Studies List


Appendix A4. Excluded Studies List


Appendix A5. U.S. Preventive Services Quality Criteria for Rating Individual Studies

Systematic Reviews

Criteria:
- Comprehensiveness of sources considered/search strategy used.
- Standard appraisal of included studies.
- Validity of conclusions.
- Recency and relevance are especially important for systematic reviews.

Definition of ratings from above criteria:
**Good:** Recent, relevant review with comprehensive sources and search strategies; explicit and relevant selection criteria; standard appraisal of included studies; and valid conclusions.
**Fair:** Recent, relevant review that is not clearly biased but lacks comprehensive sources and search strategies.
**Poor:** Outdated, irrelevant, or biased review without systematic search for studies, explicit selection criteria, or standard appraisal of studies.

Case-Control Studies

Criteria:
- Accurate ascertainment of cases
- Nonbiased selection of cases/controls with exclusion criteria applied equally to both.
- Response rate.
- Diagnostic testing procedures applied equally to each group.
- Measurement of exposure accurate and applied equally to each group.
- Appropriate attention to potential confounding variables.

Definition of ratings based on criteria above:
**Good:** Appropriate ascertainment of cases and nonbiased selection of case and control participants; exclusion criteria applied equally to cases and controls; response rate equal to or greater than 80 percent; diagnostic procedures and measurements accurate and applied equally to cases and controls; and appropriate attention to confounding variables.
**Fair:** Recent, relevant, without major apparent selection or diagnostic work-up bias but with response rate less than 80 percent or attention to some but not all important confounding variables.
**Poor:** Major selection or diagnostic work-up biases, response rates less than 50 percent, or inattention to confounding variables.

Randomized, Controlled Trials and Cohort Studies

Criteria:
- Initial assembly of comparable groups:
  - For RCTs: adequate randomization, including first concealment and whether potential confounders were distributed equally among groups.
  - For cohort studies: consideration of potential confounders with either restriction or measurement for adjustment in the analysis; consideration of inception cohorts.
- Maintenance of comparable groups (includes attrition, cross-overs, adherence, contamination).
- Important differential loss to follow-up or overall high loss to followup.
- Measurements: equal, reliable, and valid (includes masking of outcome assessment).
Appendix A5. U.S. Preventive Services Quality Criteria for Rating Individual Studies

- Clear definition of interventions.
- All important outcomes considered.
- Analysis: adjustment for potential confounders for cohort studies, or intention to treat analysis for RCTs.

Definition of ratings based on above criteria:
**Good:** Meets all criteria: Comparable groups are assembled initially and maintained throughout the study (follow-up at least 80 percent); reliable and valid measurement instruments are used and applied equally to the groups; interventions are spelled out clearly; all important outcomes are considered; and appropriate attention to confounders in analysis. In addition, for RCTs, intention to treat analysis is used.

**Fair:** Studies will be graded "fair" if any or all of the following problems occur, without the fatal flaws noted in the "poor" category below: Generally comparable groups are assembled initially but some question remains whether some (although not major) differences occurred with follow-up; measurement instruments are acceptable (although not the best) and generally applied equally; some but not all important outcomes are considered; and some but not all potential confounders are accounted for. Intention to treat analysis is done for RCTs.

**Poor:** Studies will be graded "poor" if any of the following fatal flaws exists: Groups assembled initially are not close to being comparable or maintained throughout the study; unreliable or invalid measurement instruments are used or not applied at all equally among groups (including not masking outcome assessment); and key confounders are given little or no attention. For RCTs, intention to treat analysis is lacking.

**Diagnostic Accuracy Studies**

**Criteria:**
- Screening test relevant, available for primary care, adequately described.
- Study uses a credible reference standard, performed regardless of test results.
- Reference standard interpreted independently of screening test.
- Handles indeterminate results in a reasonable manner.
- Spectrum of patients included in study.
- Sample size.
- Administration of reliable screening test.

Definition of ratings based on above criteria:
**Good:** Evaluates relevant available screening test; uses a credible reference standard; interprets reference standard independently of screening test; reliability of test assessed; has few or handles indeterminate results in a reasonable manner; includes large number (more than 100) broad-spectrum patients with and without disease.

**Fair:** Evaluates relevant available screening test; uses reasonable although not best standard; interprets reference standard independent of screening test; moderate sample size (50 to 100 subjects) and a "medium" spectrum of patients.

**Poor:** Has fatal flaw such as: Uses inappropriate reference standard; screening test improperly administered; biased ascertainment of reference standard; very small sample size or very narrow selected spectrum of patients.

Appendix A6. Reviewers of the Draft Report

Carlo Catassi, MD  
Professor of Pediatrics, Università Politecnica delle Marche, Italy

Ivor Hill, MB, ChB, MD  
Professor of Clinical Pediatrics, Section Chief Pediatric Gastroenterology, Ohio State University College of Medicine and Nationwide Children’s Hospital

Ciaran P. Kelly, MD  
Professor of Medicine, Harvard Medical School; Director, Celiac Center, Beth Israel Deaconess Medical Center

Kalle Kurppa, MD, MPH  
Tampere Centre for Child Health Research, University of Tampere and Tampere University Hospital, Finland

John Marshall, MD, MSc, FRCPC, AGAF  
Professor of Medicine, Division of Gastroenterology, McMaster University, Canada
### Appendix B1. Systematic Review of Diagnostic Accuracy Studies

<table>
<thead>
<tr>
<th>Study, year</th>
<th>Aims</th>
<th>Databases searched; Literature search dates; Other data sources</th>
<th>Eligibility criteria</th>
<th>Patients/studies</th>
<th>Characteristics of identified articles: study designs</th>
<th>Characteristics of identified articles: populations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maglione, 2016&lt;sup&gt;35&lt;/sup&gt;</td>
<td>To assess the evidence on the comparative accuracy and safety of tests used to diagnose celiac disease, including serologic tests, HLA typing, video capsule endoscopy, and endoscopic duodenal biopsy.</td>
<td>Databases: PubMed, Embase, Cochrane Library, and Web of Science. Search dates: 1990 to 2015. Additional data sources: Unpublished data were requested by the AHRQ-funded Scientific Resource Center and from manufacturers of all serologic tests.</td>
<td>Controlled trials, prospective and retrospective cohorts, case-control studies, and case series that used endoscopy with duodenal biopsy as the reference standard, applied the index test and reference standard in all subjects, enrolled a consecutive or random sample, and included ≥300 patients (unless it assessed a special population), and reported sensitivity and specificity (or data that allowed calculation).</td>
<td>56 studies and 12 prior systematic reviews (27 studies and 10 systematic reviews addressed comparative diagnostic accuracy; 23 of the studies were newly published and not included in the systematic reviews). Sample sizes ranged from 62 to &gt;12,000.</td>
<td>Systematic reviews: 10; Controlled trials: 0; Cohorts: 16; Case-control: 7.</td>
<td>1 study in US, 3 in UK, 5 in the Middle East, 1 in India, and rest in continental Europe. Race/ethnicity rarely described. All studies included symptomatic patients or those with risk factors or family history of celiac disease. 6 studies were conducted in children and/or adolescents, and an additional 3 studies included a mixed population of children and adults.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study, year</th>
<th>Characteristics of identified articles: interventions</th>
<th>Pooled results</th>
<th>Conclusion</th>
<th>Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maglione, 2016&lt;sup&gt;35&lt;/sup&gt;</td>
<td>Video capsule endoscopy: 2 systematic reviews; tTG: 3 systematic reviews and 16 original studies (3 in special populations); EMA: 2 systematic reviews and 5 original studies; DGP: 3 systematic reviews and 2 original studies; HLA typing: no evidence in general population (2 studies in special populations); Algorithms: 8 original studies</td>
<td><strong>Video capsule endoscopy</strong> &lt;br&gt; Sensitivity: 89.0% (95% CI, 82.0%-94.0%) &lt;br&gt; Specificity: 95.0% (95% CI, 89.0%-99.0%) &lt;br&gt; LR+: 12.9 (95% CI, 2.9-57.6) &lt;br&gt; LR-: 0.16 (95% CI, 0.10-0.25)</td>
<td><strong>tTG</strong> &lt;br&gt; Sensitivity: 92.8% (95% CI, 90.3%-94.8%) &lt;br&gt; Specificity: 97.9% (95% CI, 96.4%-98.8%) &lt;br&gt; LR+: 45.1 (95% CI, 25.1-75.5) &lt;br&gt; LR-: 0.07 (95% CI, 0.05-0.10)</td>
<td><strong>EMA</strong> &lt;br&gt; Sensitivity: 73.0% (95% CI, 61.0%-83.0%) &lt;br&gt; Specificity: 99.0% (95% CI, 98.0%-99.0%) &lt;br&gt; LR+: 65.6 (95% CI, 35.6-120.8) &lt;br&gt; LR-: 0.28 (95% CI, 0.17-0.41)</td>
</tr>
</tbody>
</table>

**Abbreviations**: AHRQ=Agency for Healthcare Research and Quality; CD=celiac disease; DGP=deamidated gliadin peptide; EMA=endomysial antibody; HLA=human leukocyte antigen; tTG=tissue transglutaminase; UK=United Kingdom; US=United States.
Appendix B2. Quality Assessment of Systematic Review of Diagnostic Accuracy Studies

<table>
<thead>
<tr>
<th>Study, Year</th>
<th>Search dates reported</th>
<th>Search methods reported</th>
<th>Comprehensive search</th>
<th>Inclusion criteria reported</th>
<th>Selection bias avoided</th>
<th>Validity criteria reported</th>
<th>Validity assessed appropriately</th>
<th>Methods used to combine studies reported</th>
<th>Findings combined appropriately</th>
<th>Conclusions supported by data</th>
<th>Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maglione, 2016</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Good</td>
</tr>
</tbody>
</table>
## Appendix B3. Diagnostic Accuracy Studies in Asymptomatic Populations

<table>
<thead>
<tr>
<th>Study, Year</th>
<th>Type of study</th>
<th>Screening tests</th>
<th>Reference standard</th>
<th>Setting</th>
<th>Screener</th>
<th>Age of enrollees</th>
<th>N</th>
<th>Proportion with condition</th>
<th>Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mansour, 2011</td>
<td>Cross-sectional</td>
<td>tTG IgA, tTG IgG, EMA IgA, AGA IgA, and AGA IgG</td>
<td>Biopsy</td>
<td>University Hospital Iraq</td>
<td>NR</td>
<td>Mean age, 23.4 years (range, 8 to 42 years)</td>
<td>62</td>
<td>Marsh 3 a-c: 11.3% (7/62)</td>
<td>Type 1 diabetes patients with no symptoms associated with celiac disease and no family history of celiac disease or thyroid disorders</td>
</tr>
<tr>
<td>Nevoral, 2014</td>
<td>Cross-sectional</td>
<td>tTG IgA and EMA IgA</td>
<td>Biopsy</td>
<td>Single pediatric department Czech Republic</td>
<td>NR</td>
<td>Range, 16 months-19 years</td>
<td>345 (158 asymptomatic)</td>
<td>Marsh 2 or 3: Asymptomatic, 78.5% (124/158) All children, 76% (263/345)</td>
<td>Children and adolescents examined for suspected celiac disease</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study, Year</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>AUROC</th>
<th>Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mansour, 2011</td>
<td>tTG IgA: 71% tTG IgG: 57% EMA IgA: 71% AGA IgA: 57% AGA IgG: 57%</td>
<td>tTG IgA: 93% tTG IgG: 93% EMA IgA: 96% AGA IgA: 98% AGA IgG: 98%</td>
<td>NR</td>
<td>Fair</td>
</tr>
<tr>
<td>Nevoral, 2014</td>
<td>tTG IgA &gt;10 ULN and positive EMA test: 67% <strong>Subgroups</strong> First-degree relatives (n=32): 70% Type 1 diabetes mellitus (n=40): 64%</td>
<td>tTG IgA &gt;10 ULN and positive EMA test: 83% <strong>Subgroups</strong> First-degree relatives (n=32): 81% Type 1 diabetes mellitus (n=40): 93%</td>
<td>NR</td>
<td>Fair</td>
</tr>
</tbody>
</table>

**Abbreviations:** AGA=antigliadin antibodies; AUROC=area under the receiver operating curve; EMA=endomysial antibody; IgA=immunoglobulin A; IgG=immunoglobulin G; NR=not reported; tTG=tissue transglutaminase; ULN=upper limit of normal.
<table>
<thead>
<tr>
<th>Study, year</th>
<th>Appropriate spectrum of patients</th>
<th>Adequate sample size (&gt;500)</th>
<th>Credible reference standard used</th>
<th>Reference standard applied to all patients</th>
<th>Screening test adequately described</th>
<th>Reference standard interpreted independently</th>
<th>Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mansour, 2011</td>
<td>Unclear</td>
<td>No</td>
<td>Yes; biopsy</td>
<td>Yes</td>
<td>Yes</td>
<td>Unclear</td>
<td>Fair</td>
</tr>
<tr>
<td>Nevoral, 2014</td>
<td>Unclear</td>
<td>No</td>
<td>Yes; biopsy</td>
<td>Yes</td>
<td>Yes</td>
<td>Unclear</td>
<td>Fair</td>
</tr>
<tr>
<td>Author, year</td>
<td>Study design</td>
<td>No. of centers, Country</td>
<td>Study duration Mean followup</td>
<td>Interventions</td>
<td>Patient characteristics</td>
<td>Inclusion/Exclusion criteria</td>
<td>Number screened</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------</td>
<td>-------------------------</td>
<td>-----------------------------</td>
<td>---------------</td>
<td>-------------------------</td>
<td>-----------------------------</td>
<td>----------------</td>
</tr>
<tr>
<td>Kurppa, 2014</td>
<td>RCT</td>
<td>1 center Finland</td>
<td>1 year followup</td>
<td>A. Gluten diet (n=20) B. Gluten-free diet (GFD) group (n=20)</td>
<td>A vs. B Median age (range): 42 (23-62) vs. 42 (21-74) % female: 25% vs. 45% Hypothyroidism: 10% vs. 5% Other chronic condition: 35% vs. 35% Osteoporotic fracture: 0% vs. 0% Females: Infertility or frequent miscarriages: 20% vs. 11% Median age at menarche (range): 13 (13-15) vs. 13 (9-14) years</td>
<td>Targeted screening (recruited relatives of celiac patients). Included EMA-positive adults (ages 18-75 years) who considered themselves asymptomatic (defined as an absence of: abdominal pain [&gt;3 episodes over ≥3 months interfering with function], constipation [&lt;3 bowel movements per week or difficulty during defecation], and diarrhea [≥3 loose stools/day], and extra-intestinal symptoms such as joint pain, blistering rash or unexplained neurologic symptoms, and alarm symptoms including unexplained severe weight loss, vomiting, frequent or continuous fever, or rectal bleeding). Celiac disease was defined as the presence of positive EMA and gluten-dependent enteropathy. Excluded those with a previous diagnosis of celiac disease, age &lt;18 years, evident clinical symptoms, dietary gluten restriction, severe contemporary illness or immunosuppressive medication, or ongoing or planned pregnancy.</td>
<td>Screened: 3,031 at-risk volunteers Eligible: 40 Enrolled: 40 Analyzed: 40 Withdrawals or loss to followup: None</td>
</tr>
<tr>
<td>Author, year</td>
<td>Outcomes assessed</td>
<td>Clinical health outcomes</td>
<td>Clinical health outcomes: subgroups</td>
<td>Adverse events</td>
<td>Quality rating</td>
<td>Funding source</td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>------------------</td>
<td>--------------------------</td>
<td>-------------------------------------</td>
<td>----------------</td>
<td>---------------</td>
<td>----------------</td>
<td></td>
</tr>
<tr>
<td>Kurppa, 2014</td>
<td>Serology Celiac-related genotyping, Gastrointestinal Symptoms Rating Scale (GSRS): 7-point Likert scale, higher score indicates more severe symptoms. Psychological General Well-Being (PGWB): 6-point Likert scale, higher score indicates better health-related quality of life. Short-Form 36-Item Health Survey (SF-36): 0-100, higher score indicates better subjective perception of health. Laboratory parameters: Bone mineral density, Body composition, Small bowel mucosal morphology and inflammation.</td>
<td>Gastrointestinal symptoms after 1 year, difference in mean change (95% CI): GSRS Total, -0.4 (-0.7 to -0.1); p=0.003, favors GFD GSRS Diarrhea, -0.6 (-1.1 to 0.0); p=0.052, favors GFD GSRS Indigestion, -0.7 (-1.1 to -0.2); p=0.006, favors GFD GSRS Constipation, -0.1 (-0.5 to 0.3); p=0.325 GSRS Abdominal pain, -0.2 (-0.5 to 0.2); p=0.126 GSRS Reflux, -0.5 (-0.9 to -0.1); p=0.050, favors GFD Psychological general well-being, after 1 year, difference in mean change (95% CI): PGWB Anxiety, 1.6 (0.4 to 2.8); p=0.025, favors GFD PGWB Depression, 0.3 (-0.5 to 1.2); p=0.281 PGWB Well-being, 0.5 (-1.0 to 2.0); p=0.700 PGWB Self-control, 0.3 (-0.7 to 1.4); p=0.775 PGWB General health, 0.7 (-1.0 to 2.4); p=0.532 PGWB Vitality, 0.4 (-1.5 to 2.2); p=0.670 SF-36, after 1 year, difference in mean change (95% CI): SF-36 Physical functioning, -2.8 (-8.2 to 2.6); p=0.299 SF-36 Role limitations due to physical problems, 2.3 (-12.4 to 17); p=0.749 SF-36 Role limitations due to emotional problems, 7.2 (-12.6 to 27); p=0.464 SF-36 Vitality, 6.0 (-4.3 to 16.4); p=0.245 SF-36 Mental health, 2.6 (-3.8 to 8.9); p=0.414 SF-36 Social functioning, -8.3 (-15.8 to -0.8); p=0.031, favors gluten group SF-36 Bodily pain, 0.8 (-9.8 to 11.4); p=0.881 SF-36 General health, 2.8 (-7.1 to 12.7); p=0.568 VAS: Improved in the GFD group (p=0.017) Laboratory parameters: Mean blood hemoglobin (SD), g/dL: A. Baseline: 14.3 ± 1.4, Change after 1 year: -0.2 ± 0.6 B. Baseline: 14.4 ± 1.6, Change after 1 year: -0.2 ± 0.7 Mean difference between groups, 0.0 (95% CI, -0.4 to 0.4); p=0.902 Mean serum total iron (SD), micromol/L: A. Baseline: 17.3 ± 5.7, Change after 1 year: 2.8 ± 6.8 B. Baseline: 20.0 ± 8.6, Change after 1 year: 0.3 ± 7.2 Mean difference between groups, -2.5 (95% CI, -7.0 to 2.1); p=0.288</td>
<td>NA</td>
<td>No withdrawals &quot;as a result of major symptoms or complications&quot;</td>
<td>Fair</td>
<td>Academy of Finland Research Council for Health, the Competitive Research Funding of the Pirkanmaa Hospital District, the Sigrid Juselius Foundation, the Finnish Foundation for Gastroenterological Research, the Yrjo Jahnsson Foundation, the Finnish Medical Foundation, the Foundation for Pediatric Research, and the Finnish Celiac Society.</td>
<td></td>
</tr>
</tbody>
</table>
### Appendix B5. Randomized, Controlled Trial of Treatment

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Outcomes assessed</th>
<th>Clinical health outcomes</th>
<th>Clinical health outcomes: subgroups</th>
<th>Adverse events</th>
<th>Quality rating</th>
<th>Funding source</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Body composition:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean BMI (SD), kg/m²:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A. Baseline: 26.4 ± 3.7, Change after 1 year: -0.3 ± 1.0</td>
<td>Mean difference between groups, 0.3 (95% CI, -0.5 to 1.0); p=0.451</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>B. Baseline: 27.0 ± 6.8, Change after 1 year: 0.0 ± 1.2</td>
<td>Mean % total body fat (SD):</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A. Baseline: 28.9 ± 8.2, Change after 1 year: -0.6 ± 2.4</td>
<td>Mean difference between groups, -0.5 (95% CI, -2.4 to 1.4); p=0.600</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>B. Baseline: 34.0 ± 8.9, Change after 1 year: -1.2 ± 3.4</td>
<td>Mean lumbar spine (SD), g/cm²:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A. Baseline: 1.17 ± 0.21, Change after 1 year, -0.01 ± 0.03</td>
<td>Mean difference between groups, 0.01 (95% CI, -0.01 to 0.02); p=0.338</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>B. Baseline: 1.17 ± 0.19, Change after 1 year, 0.00 ± 0.02</td>
<td>Mean femur neck (SD), g/cm²:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A. Baseline: 1.00 ± 0.12, Change after 1 year: -0.1 ± 0.03</td>
<td>Mean difference between groups, 0.01 (95% CI, -0.01 to 0.03); p=0.182</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>B. Baseline: 0.97 ± 0.14 Change after 1 year: 0.00 ± 0.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:** BMD=bone mineral density; BMI=body mass index; CI=confidence interval; EMA=endomysial antibody; GFD=gluten-free diet; GSRS=Gastrointestinal Symptoms Rating Scale; HRQOL=health-related quality of life; NA=not applicable; PGWB=Psychological General Well-Being; RCT=randomized, controlled trial; SD=standard deviation; SF-36=Short-Form 36-Item Health Survey; VAS=visual analogue scale.
# Appendix B6. Quality Assessment of Randomized, Controlled Trial of Treatment

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Randomization adequate?</th>
<th>Allocation concealment adequate?</th>
<th>Groups similar at baseline?</th>
<th>Eligibility criteria specified?</th>
<th>Outcome assessors masked?</th>
<th>Care provider masked?</th>
<th>Patient masked?</th>
<th>Attrition and withdrawals reported?</th>
<th>Loss to followup: differential/ high?</th>
<th>Analyze people in the groups in which they were randomized?</th>
<th>Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kurppa, 2014</td>
<td>Yes</td>
<td>Yes</td>
<td>Mostly</td>
<td>Yes</td>
<td>Yes</td>
<td>Not possible</td>
<td>Not possible</td>
<td>Yes</td>
<td>No/No</td>
<td>Yes</td>
<td>Fair</td>
</tr>
</tbody>
</table>