Update on celiac disease – etiology, differential diagnosis, drug targets, and management advances

Samantha A Scanlon1
Joseph A Murray1,2
1Department of Internal Medicine, 2Division of Gastroenterology and Hepatology, Mayo Clinic, Rochester, MN, USA

Abstract: Celiac disease (CD) is an immune-mediated enteropathy triggered by exposure to wheat gluten and similar proteins found in rye and barley that affects genetically susceptible persons. This immune-mediated enteropathy is characterized by villous atrophy, intraepithelial lymphocytosis, and crypt hyperplasia. Once thought a disease that largely presented with malnourished children, the wide spectrum of disease activity is now better recognized and this has resulted in a shift in the presenting symptoms of most patients with CD. New advances in testing, both serologic and endoscopic, have dramatically increased the detection and diagnosis of CD. While the gluten-free diet is still the only treatment for CD, recent investigations have explored alternative approaches, including the use of altered nonimmunogenic wheat variants, enzymatic degradation of gluten, tissue transglutaminase inhibitors, induction of tolerance, and peptides to restore integrity to intestinal tight junctions.

Keywords: immune-mediated enteropathy, gliadin, gluten, epidemiology, CD diagnosis, therapy

Epidemiology

Until the 1970s, the estimated global prevalence of celiac disease (CD) in the general population was just 0.03%.1 This estimate was based on rates of clinically detected CD and for a long time it was felt that CD was rare in many Western populations with increased frequency in places such as the west of Ireland. Subsequently it became apparent that the rate of CD diagnosis was increasing and most assumed that it was due to greater detection due to use of serology linked to heightened suspicion.2 The true prevalence is difficult to determine since there is a wide spectrum of symptoms associated with the disease.3 The iceberg model of disease has been frequently applied to CD in that the tip of the iceberg represents patients with classic malabsorption, while the more atypical presentations represent the portion of the iceberg that is invisible, but much larger, below the waterline.4 It is estimated that there may be eight times as many subclinical and silent celiac cases as there are classically symptomatic cases.5 The ratio between diagnosed and undiagnosed cases may be even greater. In a population-based study, the seroprevalence of undiagnosed CD in patients aged 50 years or older was 0.8%.6

The real prevalence of CD, based largely on serological screening in the general populations of North America and Western Europe, is estimated to be between 0.5% and 1.26%, with a somewhat higher prevalence in the Scandinavian countries, Ireland, and the UK (where it was estimated to be between 1.0% and 1.5%).3
While most studies have been based on seroprevalence, one recent study confirmed high prevalence with parallel serology and histopathology in a population-based study. These patients had no symptoms or reasons to suspect that they had CD. When including patients with positive serology, increased intraepithelial lymphocytes (IELs) without atrophy, and two previously treated patients, the prevalence was 1.8%.7

This increase in prevalence is not solely due to more effective diagnostic methods. Recent studies from the USA and Europe have shown a 2–4.5-fold increase in the prevalence of CD.5,8,9 One study showed a fourfold increase, comparing sera obtained from young men in the 1950s with present-day sera from cohorts of men with similar ages at testing and similar birth years.8 This increase in prevalence in both age cohorts suggests the presence of pervasive environmental factor(s) that are triggering the development of CD in genetically susceptible individuals of any age. A more recent study showed increased prevalence in the same cohort of adults followed over 25 years.8 There are many possible environmental factors that could be responsible, including changes in the timing of introduction, quality, quantity, or processing of cereal. In the last 40 years, there have been many changes in wheat genetics, bread processing, and enzymatic modifications of wheat prolamins. A more general explanation could be the “hygiene hypothesis,” referring to changing patterns of early childhood infections, which has also been suggested as a possible catalyst in the changing prevalence of CD4 but is unlikely to explain the new occurrence of disease in the elderly.

**Pathophysiology**

Gluten is a mixture of proteins (including gliadins and glutenins) that are found in wheat grains. Similar proteins from barley (hordeins) and rye (secalins) also induce injury.10 Of these, the gliadin peptides are the most immunogenic for CD. These peptides resist complete digestion and likely pass across the intestinal epithelial barrier via both transcellular and paracellular mechanisms. Transcellular mechanisms include IFN-γ stimulated transcytosis via intracellular vesicles.11 Paracellular mechanisms of peptide are thought to occur via increased intestinal permeability via so-called entry “leaky” tight junctions.

One of the proteins involved in controlling tight junctions is zonulin. When gliadin binds to CXCR3 chemokine receptor, zonulin is released from intestinal cells. Ex vivo studies have shown that increased levels of zonulin lead to decreased transepithelial electrical resistance (TEER). In vivo studies in mice demonstrated that zonulin increased both gastric and small intestinal permeability. The mechanism by which zonulin achieves these changes in TEER and gut permeability is via activation of proteinase-activating receptor 2 (PAR2), which in turn activates epidermal growth factor receptor (EGFR). Activation of EGFR induces an EGFR-driven decrease in TEER. Expression of zonulin mRNA is increased in individuals with active CD, suggesting that this protein is related to intestinal damage in patients with CD.12

Gliadin peptides then stimulate intestinal CD4+ T cells in the lamina propria, although there is significant variability among epitopes of gliadin in the severity of their individual immunogenic responses. The α-gliadin epitopes are recognized by T cells in nearly all patients with CD, while the γ-gliadin and glutenin epitopes are not as universally recognized.10

These epitopes on gliadin are predominantly Gln- and Pro-rich peptides that undergo deamidation of select glutamine residues by tissue transglutaminase (TTG).10,13,14 Deamidation involves conversion of neutral glutamine residues to negatively charged glutamic acid residues.15 This process results in significantly increased affinity to HLA-DQ2, and in some cases to HLA-DQ8 molecules, and leads to an enhanced antigenic presentation of gliadin.1,14,16 These HLA Class II molecules are expressed on antigen-presenting cells (ie, macrophages, dendritic cells, and B cells). Binding of deamidated peptides activates gluten-specific CD4+ T-helper 1 (Th1) cells in the lamina propria, which in turn results in intraepithelial lymphocytosis, crypt hyperplasia, production of cytokines leading to villous atrophy, and expansion of B cells that produce antibodies to gliadin and TTG.15,17 While TTG IgA and IgG antibodies are associated with the development of CD, it is not clear whether they have a pathogenic role. Anti-TTG2 IgA deposits are detected in morphologically normal jejunum before systemic detection of TTG antibody, which suggests a pathogenic role for antibodies.18

**Role of genetics in the development of CD**

The HLA molecules DQ2 and, to a lesser degree, DQ8, are intrinsic to the development of CD, playing a major role in the risk of disease development, as well as disease severity. HLA-DQ2 is found in 95% of patients with CD and is encoded by the gene pairs DQA1*05/DQB1*02xx. The alternative HLADQ2.2 molecule, encoded by DQA1*0201/DQB1*0202, is only rarely associated with CD risk though rare cases occur. Approximately 5%–10% of patients with CD carry HLA-DQ8 (encoded by HLA-DQA1*0301 and
HLA-DQB1*0302). These heterodimers are thought to be responsible for at least 40% of disease heritability. The remaining 60% is thought to be due to a number of non-HLA genes.

While HLA-DQ2 and HLA-DQ8 are found in virtually all patients with CD, it is important to note that 30%–35% of Caucasians carry these markers, of whom only 2%–5% will develop CD. This disparity and the less than 100% concordance in monozygotic twins suggest both other genetic and environmental factors are required for the development of CD. Genome-wide association studies on CD have identified non-HLA loci that are also associated with increased risk of developing CD. Many of these non-HLA loci are found in other autoimmune diseases. These loci include genes that encode for proteins such as integrins (encoded by ITGA4 at 2q31); chemokines, cytokines, and their receptors (IL2 and IL21 at 4q27, IL18RAP at 2q11–2q12, IL12A at 3q25–3q26, the CCR1 and CCR3 cluster locus at 3p21); several proteins associated with signaling pathways (RGS1 at 1q31, SH2B3 at 12q24, ATXN2 at 12q24, TNFAIP at 6q23.3, REL at 2p16.1); proteins involved in B-cell (RGS1) and T-cell activation (TAGAP at 6q25); and proteins involved in cell adhesion and motility (LPP at 3q28). Single nucleotide polymorphisms on chromosome 2q31 are also associated with CD susceptibility. Of these non-HLA loci, the 4q27 region, which hosts IL2 and IL21, shows the strongest association to the development of CD, yet it accounts for less than 1% of genetic risk. Even when combining all known non-HLA loci, they attribute for less than 10% of genetic risk.

Role of environmental factors in the development of CD

Environmental factors that contribute to the development of CD in genetically susceptible individuals include early and substantial exposure of infants to dietary gluten, early infection with enteropathic viruses, and changes in bacterial flora of the gut. The contribution of infant feeding has been debated since the 1980s and this subsequently led to the recommendation that small amounts of gluten should be gradually introduced between 4 and 7 months of age during concomitant breastfeeding. Children exposed to gluten in the first 3 months of life have a 5x increased risk of CD compared with children exposed to gluten at between 4 and 6 months of age. Children exposed to gluten after 7 months have a marginally increased risk of developing CD in comparison with those exposed at 4–6 months. The mechanism by which breast milk exerts its protective effect against the development of CD is unknown, but it is theorized that breast milk may have immune-modulating properties that suppress T cells, dilute the amount of ingested gluten, prevent gastrointestinal (GI) infections that would increase intestinal permeability and gluten exposure, or decrease gliadan uptake because IgA antibodies in breast milk agglutinate with gliadin. Infectious etiologies have been proposed to have a causal relationship with the development of CD in genetically susceptible individuals. These include adenovirus type 12, HCV, Campylobacter jejuni, Giardia lamblia, enterovirus, and rotavirus.

Socioeconomic status may also be a factor in the development of CD. An epidemiologic survey comparing schoolchildren in a prosperous neighborhood of Finland with schoolchildren living in a poorer neighborhood in Russia suggested that lower socioeconomic status may protect against CD development. It is possible, however, that children in poorer neighborhoods may have been breastfed longer for financial reasons and that the difference is related to breast milk exposure and not strictly socioeconomic status.

Clinical presentation

CD subtype

The subtypes of CD include classical, atypical, latent, and silent CD (Table 1). The classical symptoms associated with CD are diarrhea, abdominal distension, and failure to
thrive in the setting of villous atrophy.\textsuperscript{1,2} This triad is more commonly seen in individuals between 6 and 24 months of age.\textsuperscript{1,2} Atypical CD is characterized by milder GI symptoms similar to irritable bowel syndrome.\textsuperscript{1,3} It is associated with extraintestinal manifestations, such as iron deficiency anemia, osteoporosis, short stature, arthritis, infertility, peripheral neuropathy, hypertransaminasemia, and even liver failure at the time of diagnosis.\textsuperscript{1,3,17} Latent CD applies to patients who carry HLA-DQ2 and/or HLA-DQ8, with or without positive serology, and who have not yet developed villous atrophy but may have mild inflammation or immune activation may be seen.\textsuperscript{1} Patients in this subset may be asymptomatic or may have extraintestinal manifestations.\textsuperscript{1,24} Silent CD is characterized by positive serology and villous atrophy in an asymptomatic patient.\textsuperscript{3} After undertaking a gluten-free diet (GFD), some asymptomatic patients will notice improvement in different physical and psychological aspects of their life, such as improved appetite, reduced fatigue, or fewer behavioral abnormalities. In retrospect, these patients are perhaps not as asymptomatic as they thought they were prior to undertaking a GFD.\textsuperscript{1}

**Presentation of CD in children**

CD has previously been considered a disease of failure to thrive associated with classic symptoms of malabsorption: predominately weight loss, steatorrhea, and multiple deficiencies, most often presenting in very young children. However, several studies have shown that the proportion of children presenting with classic symptoms has dropped substantially over time. Indeed, most recently, the majority of patients now present with non-GI symptoms. The age at which children are presenting with CD is also increasing and is now approaches a median of 7 years.\textsuperscript{23,25}

CD often presents with failure of axial height development and delayed menarche in girls. It can often be found in family members of individuals found to have CD or in children with type I diabetes.

**Presentation of CD in adults**

When diagnosed in adults, the peak is usually in the fourth and fifth decade of life for women and men, respectively.\textsuperscript{1,26} This latter peak was previously attributed to diagnostic delay, although new studies show that CD can develop later in life in genetically susceptible individuals – presumably in response to external triggers. There has also been greater appreciation for the iceberg model of disease and, thus, increased awareness of the heterogeneity of this disease process. Increased availability of noninvasive testing, such as antiendomysial antibody (EMA), antitrautransglutaminase antibody (AGA), and TTG tests, has also increased the ability to screen patients. It is important to note that CD-specific serologic tests such as AGA and EMA did not become available until the 1990s.\textsuperscript{23}

As has been seen in children, similar changes in the frequency of GI symptoms at the time of diagnosis have been observed in the adult population. Prior to 1973, 73\% of adults presented with diarrhea; in contrast to only 43\% after 1993.\textsuperscript{27}

**Who should be considered for testing?**

Testing should be considered in all people with malabsorption, chronic diarrhea, iron deficiency anemia, vitamin D deficiency, osteoporosis, dyspepsia, and unexplained abdominal pain. Increased suspicion of CD is warranted in patients with atypical presentations of CD or who have one of the following conditions that are associated with CD: diabetes mellitus type 1, autoimmune hepatitis, autoimmune thyroiditis, vitiligo, Addison’s disease, Down’s syndrome, or dermatitis herpetiformis.\textsuperscript{23} A first-degree relative diagnosed with CD should also be considered for testing. While the inheritance of CD does not follow clear Mendelian inheritance, there is clearly a genetic predisposition to development of this disease. The prevalence of CD in first-degree relatives is estimated to be between 2.8\% and 17.2\%; in second-degree relatives, the prevalence ranges between 2.6\% and 19.5\%.\textsuperscript{3}

**Diagnosis**

Several sets of criteria recommend that the diagnosis of CD is made when there are typical small-intestinal histopathological abnormalities and a clinical response to a strict GFD. Symptoms should improve within 2 weeks. In asymptomatic patients, repeat endoscopy ensuring histopathologic resolution of disease after a GFD is undertaken would be required to fit the diagnostic criteria since these patients lack symptoms that can be used as markers of response. The presence of CD-associated antibodies (ie, TTG, anti-gliadin Ab, endomysial Ab, deamidated gliadin Ab) at the time of diagnosis and their resolution to normal after a GFD has been undertaken also support a diagnosis of CD.\textsuperscript{1,28}

**Serology**

While antigluten and other less-specific antibodies have been in use for many decades, it was only with the development of the EMA test in 1983 that the era of modern specific serology for CD was ushered in.\textsuperscript{29} In 1997, tissue transglutaminase (TTG) was identified as the actual antigen in the EMA test and became the basis for the enzyme-linked immunosorbent...
The most sensitive and specific serologic tests are EMA IgA and TTG IgA. In a systematic review of 34 studies, both EMA and TTG had high specificity (>99% and >98%, respectively), as well as high sensitivity (93% for both). Based on this study, human recombinant anti-TTG IgA is the preferred screening test.

The original antigliadin test has been supplanted by tests detecting antibodies against deamidated gliadin peptides. In a meta-analysis of eleven studies, the deamidated gliadin peptide (DGP) antibody test was compared to the TTG Ab test. The pooled sensitivities for the DGP Ab and TTG Ab tests were 87.8% and 93%, respectively. The pooled specificities for DGP Ab and TTG Ab were 94.1% and 96.5%, respectively. One of the studies in this meta-analysis selected serology from biopsy-proven patients to establish the sensitivity, specificity, and accuracy of the DGP antibody test. In this patient population, the sensitivity, specificity, and accuracy for the different tests were as follows: deamidated gliadin-IgA (74%, 95%, 86%), deamidated gliadin-IgG (65%, 98%, 84%), and deamidated gliadin-IgA+IgG (75%, 94%, 86%).

Based on this meta-analysis, deamidated gliadin antibody is a better diagnostic test for CD than gliadin antibody testing, but there is insufficient evidence to support its use over the TTG or EMA tests.

Histology

In the setting of positive serology or a high clinical suspicion despite negative serology, the next step in the diagnostic process is duodenal biopsy. Biopsy can be avoided in untreated patients with negative serology and in whom clinical suspicion is low. The spectrum of gluten-sensitive enteropathy varies from severe with total villous atrophy and crypt hyperplasia, to mild – indeed, only involving increase in inflammation in the surface epithelium – so-called intraepithelial lymphocytosis. Several schema have been developed to describe the varying degree of injury. The Oberhuber modification of the Marsh schema has been used in such a fashion (Table 2).

| Villous classification | Crypt height | Intraepithelial lymphocytosis/EC | Diagnostic of CD?
|------------------------|-------------|-------------------------------|----------------
| Type 0 | Normal architecture | Normal height | <4 IEL/100 EC | No |
| Type 1 | Normal | Normal height | >4 IEL/100 EC | No |
| Type 2 | Normal | Crypt hyperplasia | >4 IEL/100 EC | No |
| Type 3A | Mild villous flattening | Increased height | >4 IEL/100 EC | Yes |
| Type 3B | Marked villous flattening | Increased height | >4 IEL/100 EC | Yes |
| Type 3C | Total villous flattening (flat mucosa) | Increased height | >4 IEL/100 EC | Yes |

Notes: Recently this threshold was reduced to 25/100; the data displayed in this table have been collated from data in Oberhuber.

Abbreviations: CD, celiac disease; EC, enterocytes.
response. This has been described in individuals who have circulating antibodies against endomysium or TTG and often they have symptoms that will respond to a GFD.\(^1,3^4\)

While the presence of villous atrophy is specific for CD, there are other diseases also characterized by villous atrophy. A list of these diseases is included in Table 3. In the absence of other serologic and histologic markers of CD, these conditions should be considered in patients with villous atrophy.

**Interpretation of serology and histology**

Patients with positive serology and histologic changes consistent with CD are straightforward diagnoses, as demonstrated in Figure 1. However, patients with positive serology and negative biopsies represent a diagnostic dilemma, as they may represent false serologic positives, false histologic negatives, early CD,\(^2^4\) or latent CD.\(^1\) When considering the limitations of the serologic tests, it is important to note that these tests were created and conducted in populations with a higher prevalence of CD; therefore, when these tests are performed in populations with lower prevalence, the positive predictive power falls and will result in an increased number of false-positives. Conversely, the negative predictive power will increase as the test is used in populations with lower prevalence of CD; provided that the sensitivity of the test is \(\geq 50\%\). In fact, even quite specific tests such as the EMA, TTG IgA, and TTG IgG tests may be associated with as high as a \(30\%–35\%\) false-positive rate when used in low-prevalence populations.\(^5\)

**Table 3** Differential diagnosis of villous atrophy\(^3^9\)

- Celiac disease
- Tropical sprue
- Adult-onset autoimmune enteropathy
- Hypogammaglobulinemia
- Idiopathic AIDS enteropathy
- Food protein hypersensitivity
- Eosinophilic gastroenteritis
- Whipple disease
- Abetalipoproteinemia
- Intestinal lymphoma
- Collagenous sprue
- Tuberculosis
- Giardiasis
- Crohn’s disease
- Small-bowel bacterial overgrowth
- Infectious enteritis
- Parasitic infestation
- Severe malnutrition
- Small-bowel ischemia


**Abbreviation:** AIDS, acquired immune deficiency syndrome.

In patients where the clinical suspicion of CD is high but serology and/or histology (most commonly due to prior treatment with a GFD) are negative, testing for the presence of HLA-DQ2 and HLA-DQ8 may identify high-risk patients who may benefit from long-term screening.\(^1^9\) These markers carry a high negative predictive value and, therefore, if both are negative, there is a low likelihood that the patient has CD or is at risk for developing CD.\(^1\)

Some physicians have suggested that in setting of strongly positive TTG, endoscopy may not be necessary to make a diagnosis of CD. This hypothesis assumes that TTG directly correlates with histopathology and is specific enough to rule in all stages of pathology. In a study by Vivas et al, TTG titers were found to be independent predictors for Marsh III lesions in children and adults. However, it was noted that adults with CD typically experience less severe atrophy and lower TTG antibody titers than children.\(^3^2\) While strongly positive TTG titers may be sufficient to make the diagnosis of CD in children, it may be less attractive to use in diagnosing adults where the certainty provided by biopsy is easier to obtain.

The most definitive seropathological correlation was undertaken in a parallel population-based study of 1000 adults in Sweden, where all subjects had biopsies and serology with TTG. This study showed that the specificity of TTG is very high at over 98%. However, the positive predictive value in the context of a setting where the prevalence of CD was 1.6% was just 50%. Endomysial antibody positivity substantially increased this positive predictive value to virtually 100%. The study also established 25 IELs/100 enterocytes as a threshold for serologic predetermined gluten sensitivity. Even in this most rigorous setting, it is important to consider duodenal biopsy in the diagnosis of CD in adults, though a combination of positive TTG followed by a positive endomysial antibody had a combined specificity of virtually 100%.\(^7\) As such, it is not advisable to forego duodenal biopsy in the diagnosis of CD in adults since disease presentation and monitoring in adults differs greatly from children.

**Standard endoscopy**

Current standard of care is to take at least four to six biopsy specimens from the descending duodenum and bulbar mucosa.\(^4^0,4^2\) The purpose of multiple blind biopsies is to increase the diagnostic yield and avoid false negatives associated with patchy forms of CD.\(^4^2,4^4\) False negatives can also be obtained as a result of poor sampling quality, tangential sectioning, or failure to orient tissue correctly after biopsy.\(^3^3\)
A number of endoscopic signs are associated with CD, including: mosaic pattern of the duodenal mucosa, loss of duodenal folds, nodular pattern of the bulb and duodenal mucosa, scalloping of the valvulae conniventes, and visibility of submucosal vessels. These features are neither sensitive nor specific enough to be used as screening tools during endoscopy.40,42,45,46 Endoscopic clues to villous atrophy may also indicate the need for biopsies or also provide direction for targeted biopsies.41

**Enhanced endoscopy**

**Water-immersion technique**

The water-immersion technique (WIT) involves instillation of water into the duodenum that raises the duodenal villi so they appear as tiny finger-like structures. This can be used easily during regular endoscopy and the absence of villi could be then an indication for biopsy in patients who do not already have an indication for duodenal biopsy. In a prospective study of 19 children by Cammarota et al, WIT was found to have had 93% accuracy.40,44

**Chromoendoscopy**

Chromoendoscopy includes topical application of stains or pigments to improve tissue localization, characterization, or diagnosis during endoscopy. Indigo carmine staining highlights villous atrophy; however, it does not improve the ability to detect scalloping, loss of folds, or mosaic patterns.47

Altering the light use for imaging with narrow-band imaging (NBI) or Fujinon intelligence color enhancement (FICE) (Fujinon Inc, Wayne, NJ) may provide enhanced images of villi but do not consistently improve pre-CD detection in low-risk patients.48

**Video capsule endoscopy**

Video capsule endoscopy provides a method to examine the mucosa of the entire small intestine and it has been used to confirm the almost universal involvement of the proximal intestine with a widely varying extent of atrophy, which can vary from a very short length to the entire small intestine with a median involvement of 10%. The magnification and motion views enhance the detection of atrophy with a sensitivity of approximately 90%, which can be improved further with advanced image analysis techniques. Video capsule endoscopy has also been shown to heal the small intestine, but CD heals from the bottom up. It also may be able to detect complications in unresponsive CD, and it can be used to confirm villous atrophy in those patients unable or unwilling to undergo endoscopy.49 Video capsule endoscopy also allows for examination of the entire small bowel and therefore may be able to detect the presence of CD beyond the duodenum.49 Based on video capsule endoscopy studies, the mean percentage of small intestine with villous atrophy has been reported at between 16.4% and 17.9%.26,50

Our recent study has identified a quantitative method that seems to dramatically reduce the false0 positive rate of video capsule endoscopy and may be helpful in reducing interobserver variance in its interpretation. Other studies evaluating the diagnosis of CD using video capsule endoscopy

---

**Figure 1** Algorithm for the diagnosis of celiac disease.

**Note:** Nutritional deficiencies include testing iron, folic acid, calcium, vitamin D, vitamin B12 and vitamin A.

**Abbreviations:** TTG, tissue transglutaminase; EGD, esophagogastroduodenoscopy; CD, celiac disease; GFD, gluten-free diet.
have reported wide ranges in the sensitivity (70%–95.2%) and specificity (63.6%–100%). These ranges may be secondary to interobserver differences. One way to overcome this limitation involves a quantitative method that calculates mean frame-to-frame pixel brightness, image texture, and periodicity in brightness to classify and identify patients with CD. Use of these quantitative markers results in sensitivity of 92.7% and specificity of 93.5%.51

If specific complications are detected by video capsule endoscopy, plans can be made with balloon-assisted endoscopy to obtain specific samples for histologic review.

Balloon-assisted endoscopy
Balloon-assisted endoscopy is useful to exclude enteropathy-associated T-cell lymphoma (EATL), ulcerative jejunitis, and small bowel adenocarcinoma. These disease processes may be suggested by abnormalities seen during video capsule endoscopy (ie, ulcers, masses, stenosis), signs/symptoms suggestive of refractory CD, or radiographic signs of small bowel masses or of small bowel stenosis/obstruction. The antegrade approach is the preferred route of insertion of the balloon-assisted endoscopy since EATL and ulcerative jejunitis are more common in the jejunum. EATL typically appears endoscopically as multiple ulcers without the presence of a definite tumor mass. These ulcers can vary greatly and may be circumferential, discrete, or confluent ulcerative lesions that can lead to stenosis. If a mass is present, it is typically with associated necrosis or stenosis. Balloon-assisted endoscopy should also be used during follow-up of previously treated EATL to reassess the effect of surgery/chemotherapy.

Spirus enteroscopy is another technique that uses a corkscrew effect to pull the bowel over the scope to achieve deep enteroscopy. Studies on CD are awaited.

Microscale endoscopy
Confocal endomicroscopy with intravenous injection of fluorescein may be sufficient to identify villous atrophy or crypt hyperplasia – a reported sensitivity of 94% and specificity was 92%–100%. It can detect increased intraepithelial lymphocytosis.

Endocytoscopy, visualizing the mucosa with methylene blue, had a sensitivity of 88% and specificity of 100%. The following visual characteristics are significant predictors of Marsh III pathology – low number of villi per visual field, confluence of villi, irregular epithelial lining, and inability to delineate loop capillaries. While endocytoscopy can identify atrophy, it cannot identify intraepithelial lymphocytosis.

Optical coherence tomography combines ultrasound and infrared back-scattering light to visualize the mucosa and submucosa with a sensitivity and specificity of 82% and 100%, respectively. Like endocytoscopy, optical coherence tomography is limited in its inability to assess hypertrophic crypts and intraepithelial lymphocyte count.

Summary
Simple techniques can be applied by any endoscopist and should increase the detection of villous atrophy in those in whom there is a low suspicion of CD preprocedure. Use of the newer techniques is still confined to the highly specialized endoscopy centers that could further enhance our ability to detect and characterize villous atrophy.

Treatment
The GFD is the standard treatment for patients with CD and requires avoidance of wheat, rye, and barley products. This diet usually results in clinical, serologic, and histologic remission. Unfortunately, it can be a difficult lifestyle modification for patients to make given the financial burden of gluten-free products, willpower required, and high likelihood of “hidden” gluten contamination. Gluten-free foods by definition should not contain more than 20 mg/kg of gluten in total. This small amount reflects the acceptable level of contamination that may occur in food processing or the use of wheat starch from which the gluten has been removed. In a prospective study using a gluten microchallenge, as little as 50 mg per day was sufficient to cause mucosal damage after 3 months. Levels higher than this threshold may frequently contaminate foods that theoretically should be gluten free. Currently, there is little or no verification of levels of gluten contamination even in foods labeled gluten free; the regulations in the USA have been promulgated to address this important issue.

Ingestion of oats is generally safe; even ingestions of large quantities of oats in patients with CD in remission do not result in changes in histology, EMA antibody titers, or symptoms. It is important to note that antibodies to avenin, an oat prolamin, have been detected in children and adults with untreated CD. These antibodies likely represent increased food antigen exposure as demonstrated by the fact that these avenin antibodies decrease after initiation of a GFD and they are not present in patients who ingest oats after remission of their CD. Rarely, patients with CD may react to the sparse immunogenic peptide motifs in oats. Commercial varieties of oats are often contaminated with wheat or other gluten-containing grains; therefore,
especially grown and tested varieties are necessary for the diet.

Micronutrient deficiencies are common. It is equally important to consider that the adoption of a GFD may result in a nutritionally unbalanced diet with poor intake of valuable micronutrients. Vitamin B supplementation with B12, folate, and B6 may result in normalized plasma total homocysteine (a marker of vitamin B status) and improvements in well-being – specifically, reduced anxiety and depression. Zinc and copper deficiencies can occur in CD and result in significant sequela.62

In patients where the diagnosis is not certain (ie, patients with positive serology and nondiagnostic biopsies), it is unclear whether these patients should adopt a GFD. In a study by Kurppa et al, patients with positive EMA serology, HLA-type requisiste, and either Marsh I or Marsh II were randomized to gluten-free and unrestricted diets.24 After 1 year, EMA titers, intraepithelial lymphocytosis, small bowel mucosal morphology, and symptoms improved in patients on a GFD, whereas those who continued an unrestricted diet had no improvement in these markers of disease activity and experienced progression of small bowel mucosal morphology. They otherwise did not experience worsening serologic markers.24 Some may argue that it is not prudent to wait for villous atrophy to develop since there may be complications that may not be reversible (ie, osteoporosis, dental enamel defects, and growth retardation).

While it has been held in the past that there was little evidence to support the detection of or the treatment of asymptomatic CD, recent work from several centers has suggested that patients with supposedly asymptomatic disease are not asymptomatic after they have been, in retrospect, quite symptomatic on follow-up. In addition, asymptomatic disease in diagnosed CD may be associated with an increased mortality, though this theory remains controversial.4 There is little doubt that even asymptomatic disease can be associated with significant consequences; for example, for bone health there are conflicting reports on the benefits that a GFD can affect fracture rate, though bone density may improve in many patients on a strict GFD.63 There is probably less benefit to be obtained from screening and identifying CD in elderly individuals who have no symptoms as their long-term potential for developing symptoms may be limited because of their age, whereas younger individuals would have the most benefit from having CD identified and treated.6 With respect to fracture risk, there is conflicting evidence about whether a GFD in an asymptomatic patient with CD would abrogate increased fracture risk.2 This is an important point to note, since fracture risk is frequently given as one of the key reasons for undertaking a strict GFD.

Adherence to a gluten-free diet: assessing compliance and recognizing obstacles

Lifelong adherence to a GFD is central to the treatment of this disease. Most patients will respond symptomatically to a reduction or elimination of gluten from their diet.64 While strict adherence to a GFD may be achievable for some, many patients have incomplete adherence.65 Several factors can affect adherence – appropriate education and counseling, family and social support, and psychological factors. Contamination of the GFD may occur because of deliberate and intentional ingestion of gluten, insufficient information of gluten content of foods, and contamination in the processing and preparation of foods. Food prepared in restaurants or outside the home provides the opportunity for gluten contamination and may be beyond the patient’s direct control.

Adherence to a GFD can be assessed by a face-to-face consultation with a registered dietician with expertise in the GFD and by serology.66 Typically, 6 months after the diagnosis the serologic levels should have dropped significantly and should continue to show a downward trajectory and by 1–2 years after exclusion from the diet, serology should become negative. However, a negative serology does not necessarily guarantee that the small intestine has healed.67

Alternatively, a positive serology, suggests that intestinal damage will have persisted and that there is likely gluten contamination in the diet. Finally, follow-up biopsies often reveal failure to heal, and the most common reason for this is likely some degree of gluten contamination.68

Novel therapeutic approaches

A patient who adheres to a GFD could easily ingest a significant amount of gluten beyond safe thresholds, given the allowance of 20 mg/kg of gluten in “gluten-free food.” New treatment modalities focus on modification of dietary food products, enzymatic degradation or binding of gluten to decrease the degree of gluten exposure, inhibition of small intestinal permeability, blocking the deamination of gluten via inhibition of TTG, and modulation of the immune system.1,69,71

Modification of dietary food products

Consideration has been given to “detoxifying” wheat to remove or scramble the toxic peptides of gluten. This approach is problematic since wheat gluten proteins account
for the continuous viscoelastic network in dough, thereby determining the dough’s strength and baking properties. Similar challenges have been found when attempting to select grains with low or absent immunogenic sequences.69

**Enzymatic degradation of gluten: the “glutenase” approach**

Endopeptidases (aka glutenases) aim to degrade gluten by targeting the proline-rich peptides which otherwise resist the body’s natural proteases and contain highly immunogenic peptides. This degradation would hopefully also destroy the epitopes, thereby preventing gluten’s usual immunogenic response in patients with CD.10,70 These glutenases are designed to serve as adjuvant therapy to prevent harm from incidental low-level gluten exposure during a GFD.70

Prolyl endopeptidases (PEPs) are glutenases that are ingested with a meal and become activated in the upper intestine where they work with pancreatic and brush border membrane proteases.10,14 Gluten is composed of many different gliadin and glutenin polypeptides. Approximately 60% of the gliadin proteins are α-gliadins, 30% are γ-gliadins, and 10% are ω-gliadins. An in vitro study has demonstrated that treatment of gluten with PEPs results in degradation of gluten to smaller peptides which decrease T-cell activation and replication in patients with CD.10 In an in vitro study by Marti et al, *Flavobacterium meningosepticum* PEP treatment led to elimination of all α-gliadin epitope–bearing peptides and some decrease of γ-gliadins epitopes.10 In an in vivo double-blind crossover study by Pyle et al, patients who developed fat or carbohydrate malabsorption after a 2-week gluten challenge experienced decreased malabsorption after administration of PEP.14

ALV003 is a combination of two glutenases – a cysteine endoprotease derived from germinating barley seeds in addition to a PEP from *Sphingomonas capsulate*. In a study by Tye-Din et al, ALV003 was given to HLA-DQ2 CD patients as part of a gluten challenge study.70 ALV003 was effective in decreasing gluten-specific T-cell responses, as identified via enzyme-linked immunospot assay (ELISpot). Further studies are underway to study the effect on clinically-relevant outcomes.

Some limitations of endopeptidase therapy are that it requires a sufficient amount of enzyme to be ingested and also requires sufficient exposure of gliadin to the enzyme to ensure adequate degradation.21 Copolymer binders of gluten sequences have also been studied experimentally as a potential means of treating CD. Studies in gluten-sensitized mice have demonstrated that this can effectively reduce the immune response consequently as gluten exposure.71 This has not yet been studied in humans.

**Inhibition of small intestinal permeability**

Increased intestinal permeability is a constant and perhaps early event in CD. Zonulin has been identified as a key mediator of gluten-induced permeability changes and is increased in CD. This may provide an ideal drug target. Increased serum concentrations of zonulin have been detected in patients with CD.31 AT-1001 is an octapeptide inhibitor of paracellular permeability that inhibits gliadin-induced cytoskeleton rearrangement of intestinal epithelial cells, tight junction disassembly, and peak F-actin increment. So far, it appears that larazotide is safe, generally well tolerated by patients, and prevents increased intestinal permeability. Further studies are underway to better define the effects of this new drug.72

**TTG inhibitors**

Inhibition of TTG is an attractive therapeutic possibility since TTG deamidates gluten, thereby increasing gluten’s immunogenic response, in addition to acting as a proinflammatory agent itself. TTG is involved in many other biologic pathways, such as apoptosis, cell adhesion, signal transduction, collage assembly, and wound repair mechanism. Given its broad range of actions, universal inhibition of this molecule could have disastrous results and therefore may limit its role as a drug target.69 In mouse studies, a dihydroisoazole derivative has shown promise as a TTG inhibitor; it has good oral bioavailability, a short serum half-life, and has shown effective TTG inhibition in small intestinal tissue. Mice tolerated the compound for 2 weeks, although there was some suggestion that this compound has the potential for hepatotoxicity.73

**Immune system modulation**

Under the hypothesis that the disappearance of intestinal parasites from humans in developed countries has created a situation where some individuals develop an autoimmune disease, some investigators are exploring the role of parasites in the treatment of inflammatory bowel disease and CD. A randomized double-blind placebo-controlled study evaluating immunity and disease activity of patients with CD before and after inoculation with the human hookworm *Necator americanus* was recently completed. The investigators surmise that reintroduction of this parasite will lead to interference with the intestinal immune system, which will allow for decreased gluten sensitivity.74
Desensitization therapy with peptide-based therapeutic vaccination

Nexvax2 (Nexpep, Ivanhoe, Victoria, Australia) is a peptide-based therapeutic vaccine that has induced immune tolerance in rodent models of HLA-DQ2-restricted T-cell immunity to gluten. Its safety is being currently investigated in patients with treated CD.75

Nonresponsive CD

As many as 10%–20% of patients diagnosed with CD will have persistent or recurring symptoms despite adherence to a GFD76 and there may be many reasons for this. Most commonly, it is due to contamination of the diet with gluten. It is also important to review the original diagnosis, as the diagnosis of CD may have been incorrect, resulting in failure to respond to a GFD. Additional diagnoses occurring in association with CD are quite common and include: lactose intolerance, microscopic colitis, small intestinal bacterial overgrowth, pancreatic exocrine insufficiency, and functional disorders of the gut.77

Complications, most significantly of EATL, may be preceded by refractory CD (RCD) with evidence of clonality detectable on biopsy. RCD itself is relatively uncommon, even among patients with nonresponsive disease.78 Inflammation can lead to stricturing in the intestine and substantial motility disorders can also occur and cause persistent or recurring symptoms.79

The leading cause of nonresponsive CD is persistent gluten exposure and this is seen in 30% of patients who report persistent symptoms. To assess adherence, current practice is to examine the patient’s TTG titers and to have the patient evaluated by an expert dietician. While TTG has been validated to serve as a marker of diagnosis, its utility as a marker of adherence has not been validated. Another dilemma with this practice is that many patients do not have access to an expert registered dietician. A seven-question standardized questionnaire, developed by Leffler et al, incorporates the presence of CD-related symptoms, self-efficacy, reasons for keeping to a GFD, and perceived adherence to a GFD. This questionnaire has a sensitivity of 74%, specificity of 77%, positive predictive value of 50%, and a negative predictive value of 90% when compared with the methods employed by an experienced dietician. This tool could be used as a screening method to identify patients who would benefit from a session with an expert dietician.66

If gluten exposure is resolved and symptoms persist, other etiologies to consider include: RCD, intestinal T-cell lymphoma, small intestinal bacterial overgrowth, disaccharide intolerance, irritable bowel syndrome,80 tropical sprue, collagenous sprue, adult-onset autoimmune enteropathy, hypogammaglobulinemia, and Crohn’s disease.39 Figure 2 demonstrates an algorithm that can be used as a guide to evaluate the patient with nonresponsive CD.

Figure 2 Nonresponsive celiac disease algorithm.

Abbreviations: CD, celiac disease; GFD, gluten-free object; HLA, human leukocyte antigen; CC, collagenous colitis; LC, lymphocytic colitis; EGD, esophagogastroduodenoscopy; TCR, T-cell receptor; UJ, ulcerative jejunitis; EATL, enteropathy associated T-cell lymphoma; IBS, irritable bowel syndrome.
RCD

RCD is defined by persistence or recurrence of clinical and histological symptoms despite strict adherence to a GFD for at least 6–12 months in the absence of other causes of nonresponsive CD.1,39 The incidence per 100,000 person-years was 0.06 according to the US Caucasian population in 2000.39 It is more common in patients with adult-onset CD, particularly those diagnosed above the age of 50 years,1 and is rarely diagnosed in patients under the age of 30.39

RCD should be considered in patients with either new or persistent diarrhea, abdominal pain, involuntary weight loss, multiple vitamin deficiencies, anemia, fatigue, or malaise. Most patients with RCD have normal serology, however, others may have persistently positive serology. Reasons for persistent positive serology include CD-specific antibodies kinetics, upregulation of TTG in response to severe inflammation or mucosal destruction, as well as coexisting autoimmune diseases associated with false-positive CD serology.39

The condition is divided into RCD type I and type II. Type I is characterized by an increased but phenotypically normal intraepithelial T lymphocytes in the intestinal mucosa.1,39,81 Type II is characterized by phenotypically aberrant intraepithelial T lymphocytes in the intestinal mucosa. Detection of such abnormal lymphocytes is performed on either fixed (immunohistochemistry and T-cell receptor clonal rearrangement by polymerase chain reaction [PCR]) or fresh frozen intestinal tissue (flow cytometry). Abnormal lymphocytes are classified based on: (1) loss of normal surface markers (CD3, CD4, CD8) is <50% of those with expression of intracytoplasmic CD3 in >50% of intraepithelial lymphocytes when evaluating with immunohistochemistry, or >20% if using flow cytometry in addition, and (2) T-cell receptor chain clonal rearrangements, as detected by PCR.39

Type I and type II RCD have significantly different prognoses; type I RCD is typically associated with a 5-year survival of 80%–96%, while type II is associated with a 44%–58% 5-year survival.1,82 Five factors are associated with increased mortality if present at the time of diagnosis of RCD. These include: albumin ≤ 3.2 g/dL, hemoglobin ≤ 11 g/dL, age ≥ 65 years, the presence of aberrant IELs, and total villous atrophy.82

The difference in prognosis is related to the fact that RCD type II has a higher risk of developing lymphoma arising from the clonal expansion of aberrant lymphocytes and ultimate transformation into EATL.1,81,82 EATL is characterized by nonmonomorphic cytology, CD56 negativety, and chromosomal gains of 1q and 5q.39 EATL will develop within 5 years in 60%–80% of patients diagnosed with RCD type II.81 Despite aggressive treatment, the 5-year survival rate for EATL is 8%–20%.39 A new, more intensive, regimen of chemotherapy has been shown to increase the survival rate, but the performance status of many patients may be insufficient to prevent such an aggressive regimen.83 EATL should be considered in patients with CD who present with recurrence of malabsorption, fever, abdominal pain, weight loss, and night sweats. EATL can occur throughout the GI tract, lungs, ribs, and spleen. When it occurs within the GI tract, it has a predilection for the proximal jejunum.1

Patients with type II are also more likely to have ulcerative jejunitis, large ulcerations (>1 cm), cavitating mesenteric lymph node syndrome, small splenic volume (<122 cm⁴), intussusception, bowel wall thickening, and abdominal lymphadenopathy.38

Treatment options for both types of RCD are limited and are based on small observational experiences, case reports, and expert opinion. The paucity of studies is due to it being a rare disease; therefore, it is difficult to recruit enough patients to undergo a randomized controlled trial. Treatment options for RCD type I include prednisone, budesonide, and combination therapy of azathioprine and prednisone.1,39 Patients with RCD type II have a variable response to azathioprine, prednisone, cyclosporine, infliximab, alemtuzumab, and IL-10 therapy.39,81 New therapies for RCD type II are being investigated and include interleukin-15-blocking antibodies, cladribine, and autologous stem cell transplant (ASCT).81 In a small study of seven patients by Al-toma et al, patients with RCD type II who underwent ASCT had a significant reduction in the aberrant T cells in duodenal biopsies and clinical response.81 This histologic finding was coupled with improvement in well-being as well as normalization of hematologic and biochemical markers.75

The degree of diarrhea, weight loss, and malnutrition in patients with RCD may become severe enough to require total parenteral nutrition. Close monitoring and treatment of trace mineral deficiencies (ie, copper and zinc) and bone density should also be performed. While there is no evidence to show that there is benefit to continuing a GFD in patients with RCD, strict adherence to a GFD is still widely recommended in all patients with RCD.39

Follow-up

Asymptomatic patients may have persistent villous atrophy and therefore carry substantial risk of the complications of untreated CD. Biopsy remains the gold standard for assessing histologic response to a GFD. Follow-up practice at the authors’ institution is repeat serology in 6 months and follow-up endoscopy in 1–2.5 years.
The purpose of following histology after initiation of a GFD is to assess for mucosal recovery; patients with persistent mucosal damage may have a higher risk of all-cause mortality in comparison with patients whose mucosa returns to normal. These patients with persistent mucosal damage despite resolution of symptoms may present a "latent form of RCD" given their increased mortality.

In a cohort of 241 adults who underwent diagnostic and follow-up biopsy, the rate of confirmed mucosal recovery at 2 years following diagnosis was 34% and was only 66% after 5 years. This is in contrast with recovery in children, where up to 95% of patients diagnosed during childhood will experience complete mucosal recovery within 2 years of starting a GFD. Clinical response was documented in 82% of patients; however it was not a reliable marker of recovery. There was a trend associated with mucosal recovery and reduced rate of all-cause mortality, although statistical significance was not achieved.

**Prognosis**

Clinically diagnosed but untreated CD can result in substantial morbidity and mortality, because of complications including malabsorption, increased risk of malignancies, and, in young individuals, suicide. The excess mortality in those who have been clinically diagnosed is probably modest and varies between a standardized mortality ratio (1.2–2) when all ages are included.

However, the hazard ratio for mortality in some younger individuals may actually be higher than that; in one Swedish study it was > 6x the expected rate. However, thankfully, the absolute rate of mortality is still low. As a matter of concern, individuals diagnosed with CD at a young age appear to have an increased risk of death by suicide.

While mortality estimates in diagnosed and presumably treated CD are reasonably consistent, that is not the case for undiagnosed and untreated CD. In a study of young adults (mean age 20.5 years) with undiagnosed CD, there was nearly a fourfold increased risk of death during a 45-year follow-up. A study from Finland examined patients (mean age 50 years) with undiagnosed CD but did not find such a dramatically increased risk of death; however, this population still had an increased risk of lymphoma and esophageal cancer. A recent nested case-control study from the USA compared serologically defined subjects with CD with age- and sex-matched sero-negative controls (all patients were age ≥ 50 years) and reported that this population of patients were more likely to have osteoporosis, hypothyroidism, lower body mass index, lower cholesterol, and lower ferritin. They did not appear to have any increased mortality compared with the control subjects.

It appears that there may also be an age-dependent excess in mortality, though it is also possible that decades of untreated CD are necessary for a substantial excess of mortality to occur. Further studies encompassing many decades of follow-up in individuals where CD has been retrospectively established are unlikely to be translated to a prospective, randomized evaluation of intervention with a GFD of sufficient duration to capture an effect on survival. Hence, it is the young individuals in whom CD can be detected long before consequences have occurred who are probably most likely to develop negative consequences of disease if followed for decades. Individuals of an advanced stage may have insufficient life expectancy to accrue such an increased risk. Complex issues of competing mortality and potential benefits of undiagnosed CD in the face of other population risks, such as obesity and hypercholesterolemia, will need to be considered.

**Conclusion**

While there have been a number of advances in the understanding of the genetic susceptibility, pathogenesis, and adverse outcomes associated with CD, there are still many avenues of investigation. Further advances in diagnosis are needed to identify sufficiently specific and sensitive serologic markers so that the need for endoscopy may be eliminated. In the interim, advances in endoscopy will continue to find new ways to examine in vivo histology, leading to increased diagnostic yield of biopsies, and potentially eliminating the need for biopsies. While there are no promising cures at this time, there are a number of potential therapies that may help protect against accidental gluten ingestion and its subsequent immunogenic response.

**Acknowledgment**

Joseph A Murray was supported in part by NIH grants DK 71003 and 57892.

**Disclosure**

JAM has been a consultant to Alvine Inc, Shire, ImmunosanT Inc, ActoGeniX NV, Artielle Immunotherapeutics, Ironwood Pharmaceuticals, Inc, Ferring Research Institute, Inc, Bayer Healthcare Pharmaceuticals. SAS has no conflicts of interest to declare in this work.

**References**

