Analytical Framework

The analytical framework is presented in Figure 1. In this framework, we wanted to represent the diagnostic pathways and the potential outcomes of testing various populations for CD. Each step of the pathway represents a portion of this systematic review, starting with the identification of the populations of interest, their diagnostic pathways, and ultimately the clinical outcomes, as well as consequences of testing.

Figure 1: Analytic framework

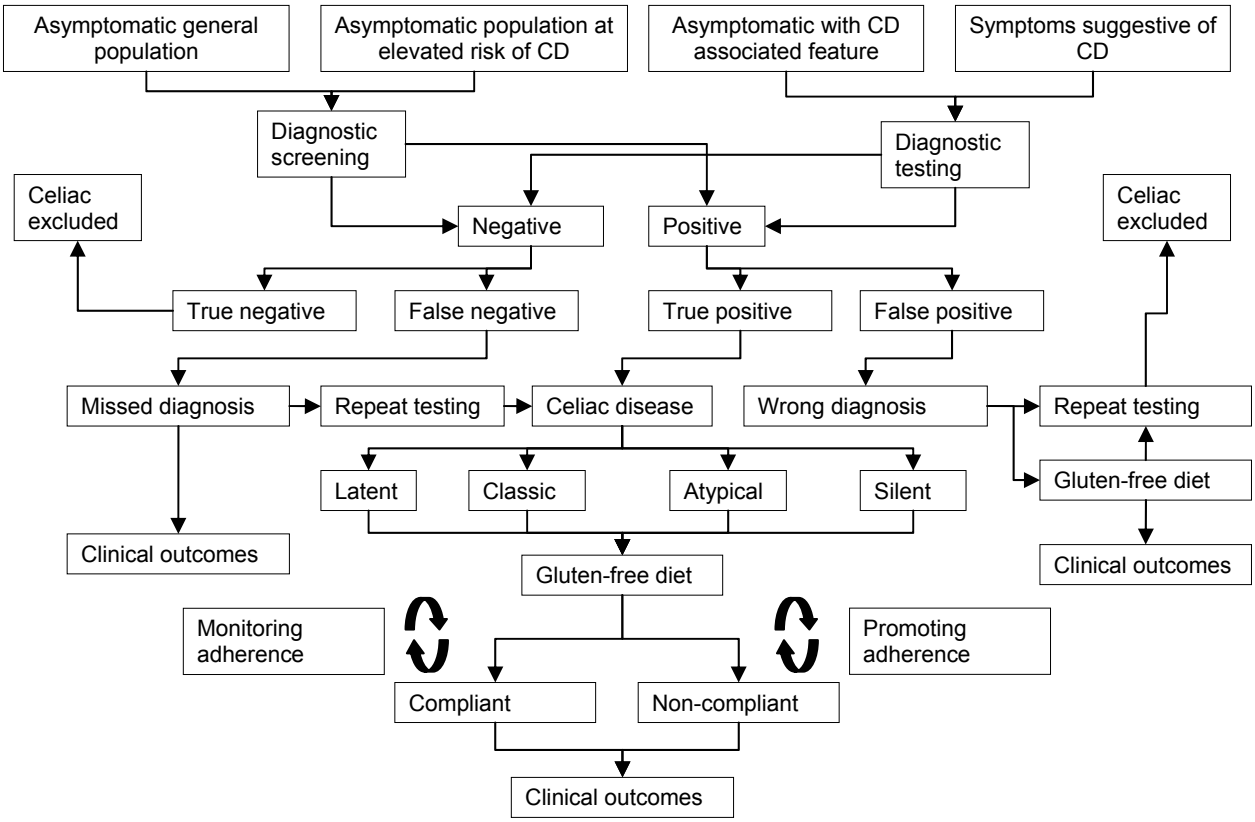


Table 4: Included studies for IgA-AGA in studies including both children and adults

Author, year; country	Study type	Biopsy criteria	Notes	Sens	Spec	PPV	NPV	Prev
Cataldo, 2000; Italy	Case-control	Original & revised criteria?	20 IgA-deficient CD vs healthy IgA-deficient non-CD	0	100	0	33.3	0.7
Sulkanen, 1998; Finland	Case-control	ESPGAN		84.5	81.6	75.2	89	0.4
Ascher, 1996; Sweden	Relevant clinical population	ESPGAN		90.9	98.5	98	92.7	0.5
Carroccio, 2002; Italy	Relevant clinical population	Marsh, broken down by criteria; CD was diagnosed as enlarged crypts and/or villous atrophy-with normalization on GFD		67	90	86	75	0.5
Tesei, 2003; Argentina	Relevant clinical population	Marsh II to IV - with confirmation		64	92	92	64	0.6

Figure 2: IgA-AGA in children with CD

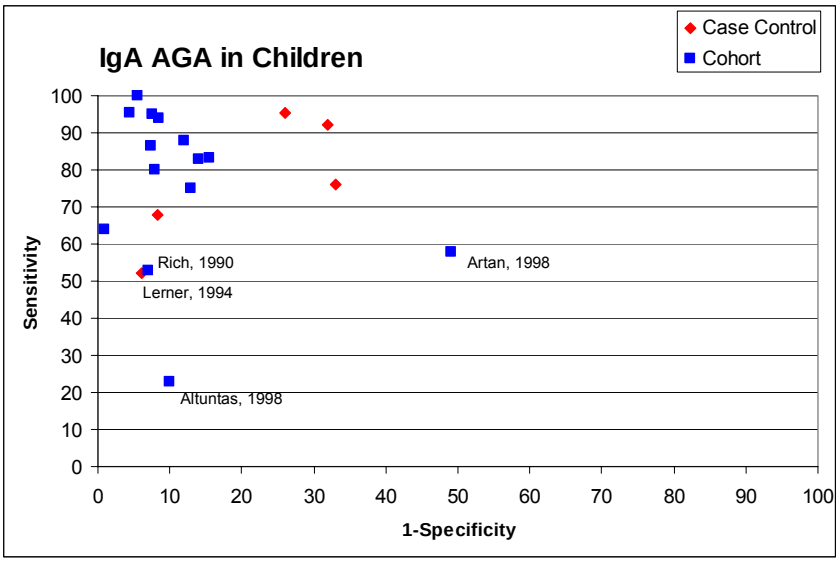


Figure 3: IgA-AGA in adults with CD

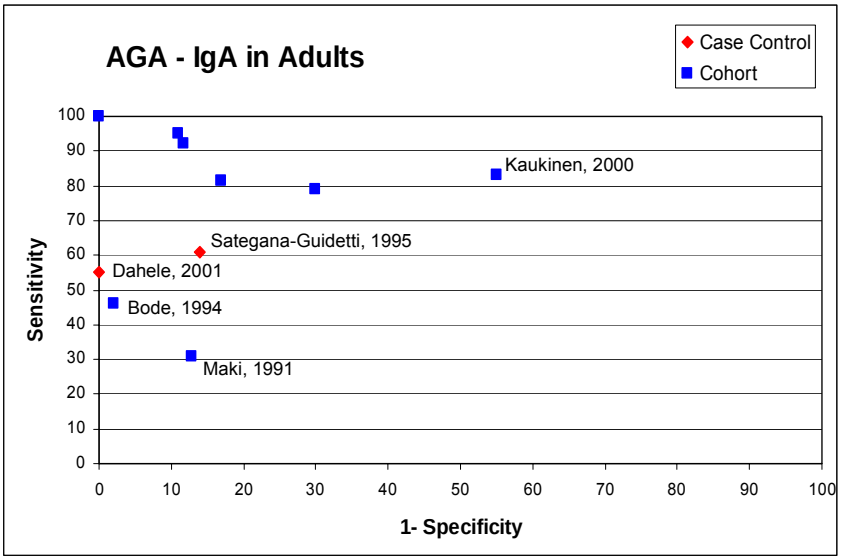
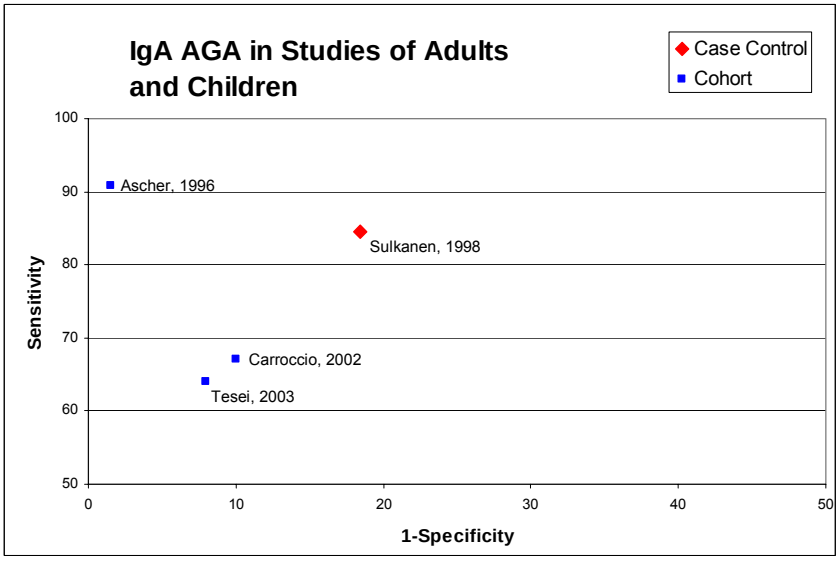


Figure 4: IgA-AGA in adults and children with CD



IgG-AGA. The seven studies of IgG-AGA in adults demonstrated considerably greater heterogeneity.[30,33,54,62,63,71,80] The sensitivity ranged from 17% to 100%, with little study grouping. However, there was less variation in the reported specificities. Five of the seven studies demonstrated specificities greater than 80%, whereas, the remaining two studies had specificities of greater than 70%. (Table 5; Figure 5)

Table 7: Included studies for IgG-AGA in studies including both children and adults

Author, year; country	Study type	Biopsy criteria	Notes	Sens	Spec	PPV	NPV	Prev
Cataldo, 2000; Italy	Case-control	Original and revised criteria?	20 IgA-deficient CD vs healthy IgA-deficient non-CD	100	100	100	100	0.7
Sulkanen, 1998; Finland	Case-control	ESPGAN		69	73.4	63	78.3	0.4
Ascher, 1996; Sweden	Relevant clinical population	ESPGAN		96.4	69.2	72.6	95.7	0.5
Carroccio, 2002; Italy	Relevant clinical population	Marsh-broke down by criteria; CD was diagnosed as enlarged crypts and/or villous atrophy - with normalization on GFD		76	75	73.4	77.3	0.5
Tesei, 2003; Argentina	Relevant clinical population	Marsh II to IV - with confirmation		84	86	89	79	0.6

Figure 5: IgG-AGA in adults with CD

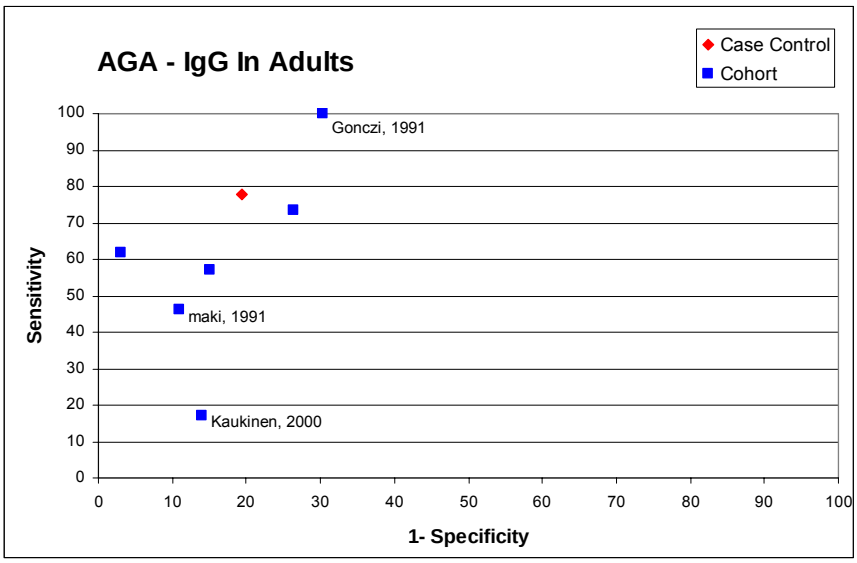


Figure 6: IgG-AGA in children with CD

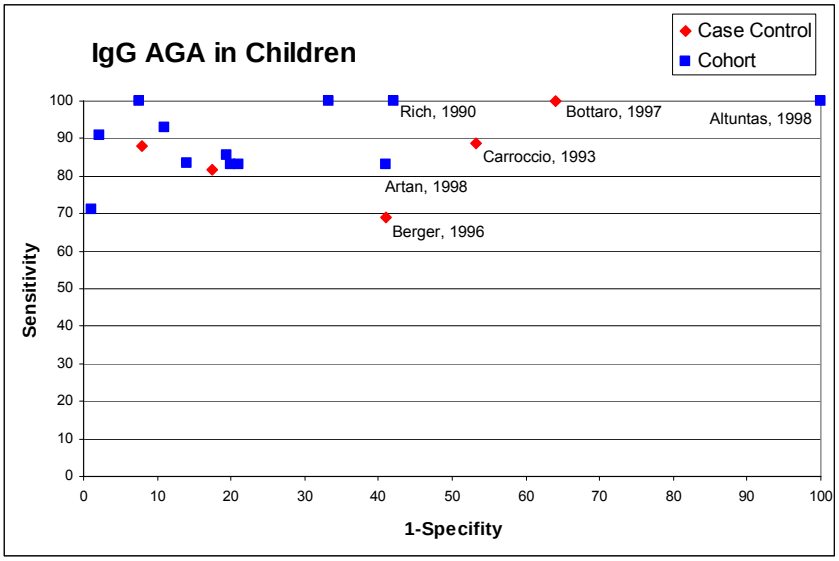


Figure 7: IgG-AGA in children and adults with CD

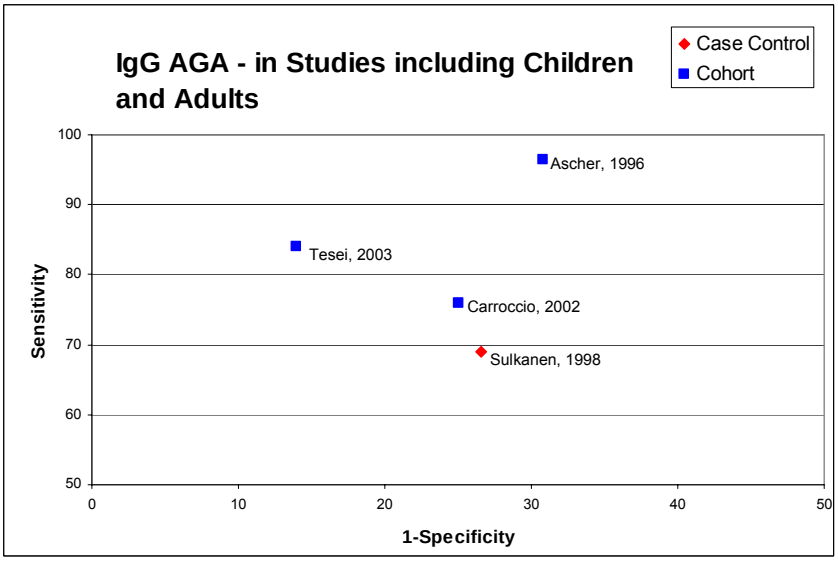


Table 10: Included studies for IgA-EMA-ME in studies including both children and adults

Author, year; country	Study type	Biopsy criteria	Notes	Sens	Spec	PPV	NPV	Prev
Cataldo, 2000; Italy	Case-control	Original & revised criteria?	20 IgA-deficient CD vs healthy IgA-deficient non-CD	0	100	0	33.3	0.7
Dickey, 2001; Northern Ireland	Case-control	Villous atrophy		75.3	98.3	98.2	76	0.6
Ascher, 1996; Sweden	Relevant clinical population	ESPGAN		98.2	100	100	98.5	0.5
Carroccio 2002; Italy	Relevant clinical population	Marsh - broke down by criteria; CD was diagnosed as enlarged crypts and/or villous atrophy - with normalization on a GFD		88	99	98.7	90	0.5
Tesei, 2003; Argentina	Relevant clinical population	Marsh II to IV - with confirmation		86	100	100	83	0.6

Figure 8: IgA-EMA-ME in adults with CD

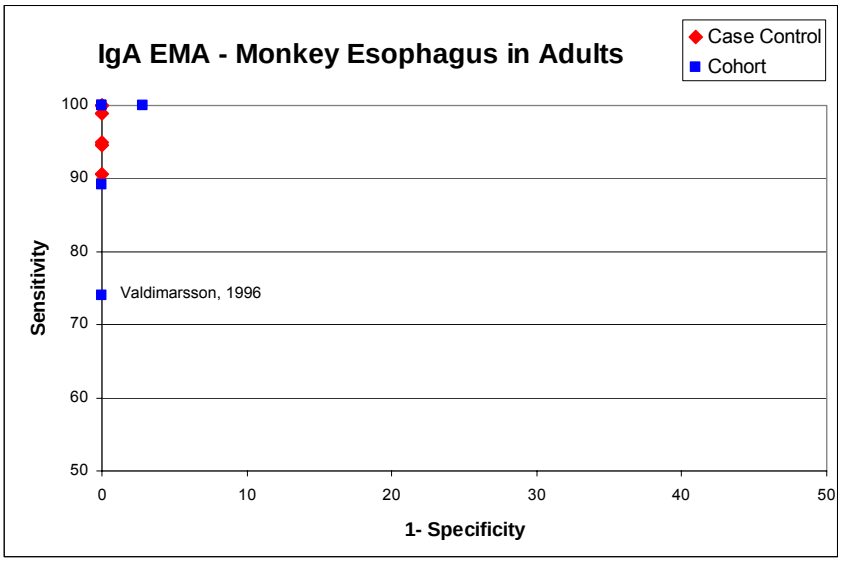


Figure 9: IgA-EMA-ME in children with CD

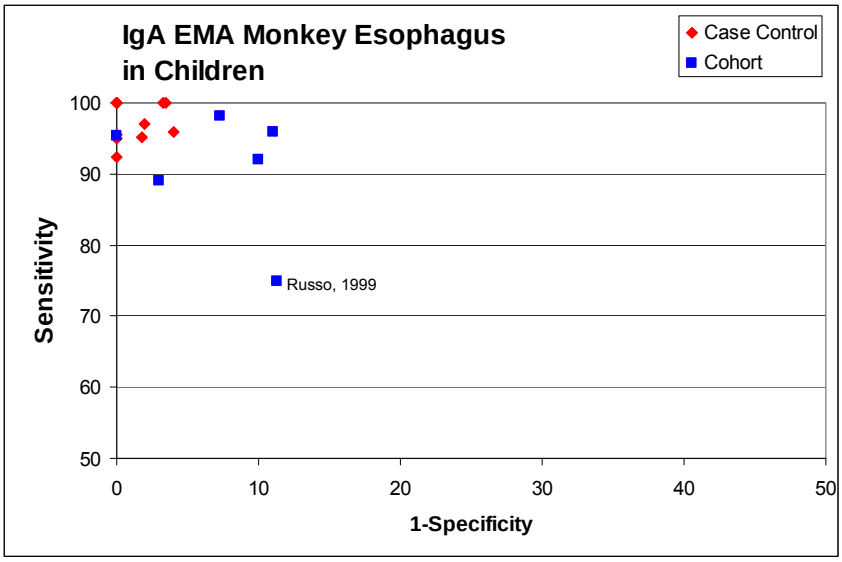
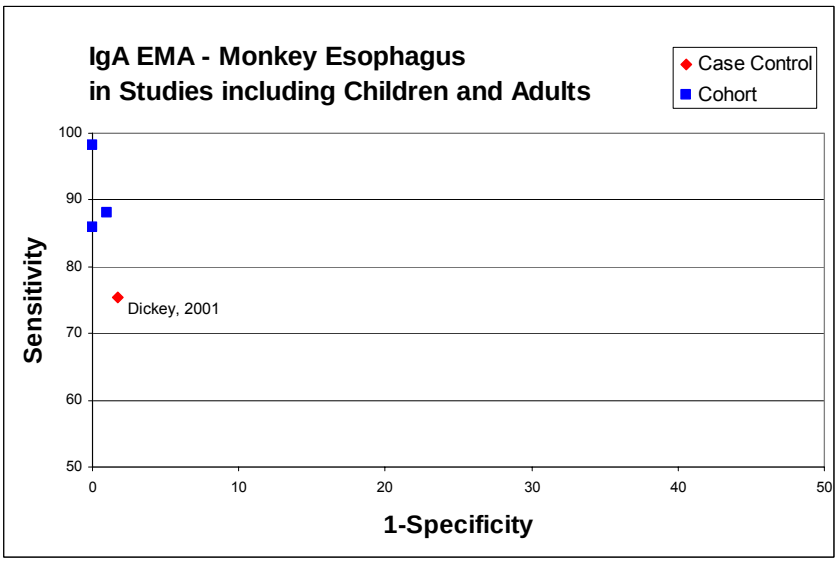


Figure 10: IgA-EMA-ME in adults and children with CD



IgG-EMA-ME. Only two studies meeting our inclusion criteria assessed IgG-EMA-ME, one in adults (Table 11),[63] and one in children (Table 12).[66] In the single adult study,[63] the sensitivity of the test was found to be 39%, whereas, the specificity was 98%. In a case-control study design, Picarelli et al. studied 30 IgA-EMA-negative children suspected of having CD.[66] Of these 30 children, 18 were subsequently found to have CD by duodenal biopsy and nine of the 18 were

Table 15: Included studies for IgA-EMA-HU in studies including both children and adults

Author, year; country	Study type	Biopsy criteria	Sens	Spec	PPV	NPV	Prev
Sblaterro, 2000; Italy	Case-control	ESPGAN	93	100	100	80	0.8
Sulkanen, 1998; Finland	Case-control	ESPGAN	92.6	99.5	99.2	94.9	0.4

Figure 11: IgA-EMA-HU in adults with CD

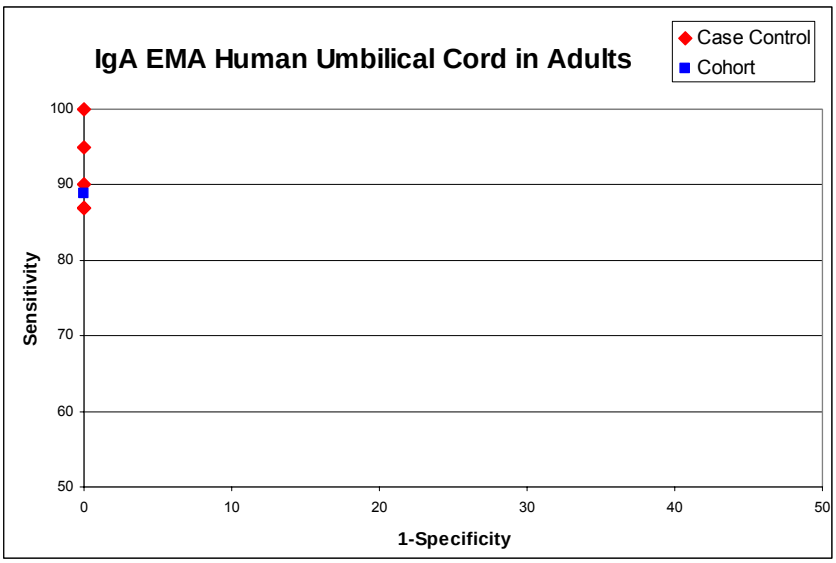


Figure 12: IgA-EMA-HU in children with CD

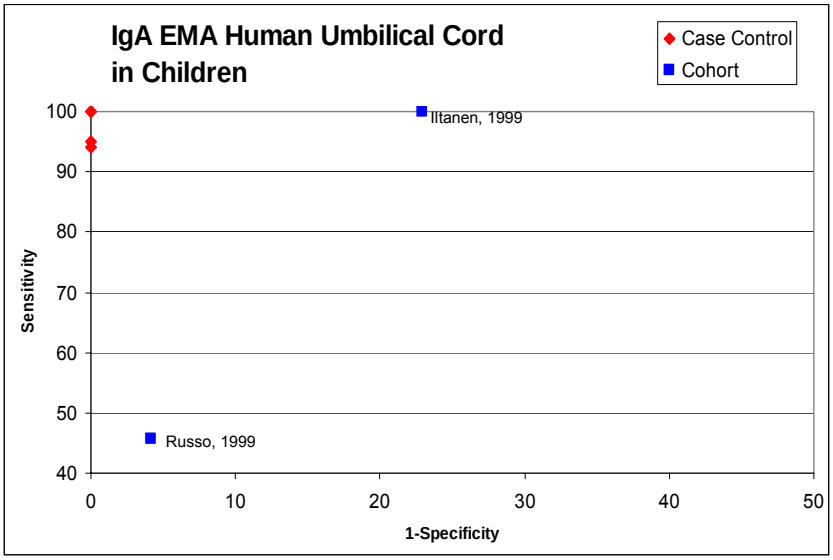


Figure 13: IgA-EMA-HU in adults and children with CD

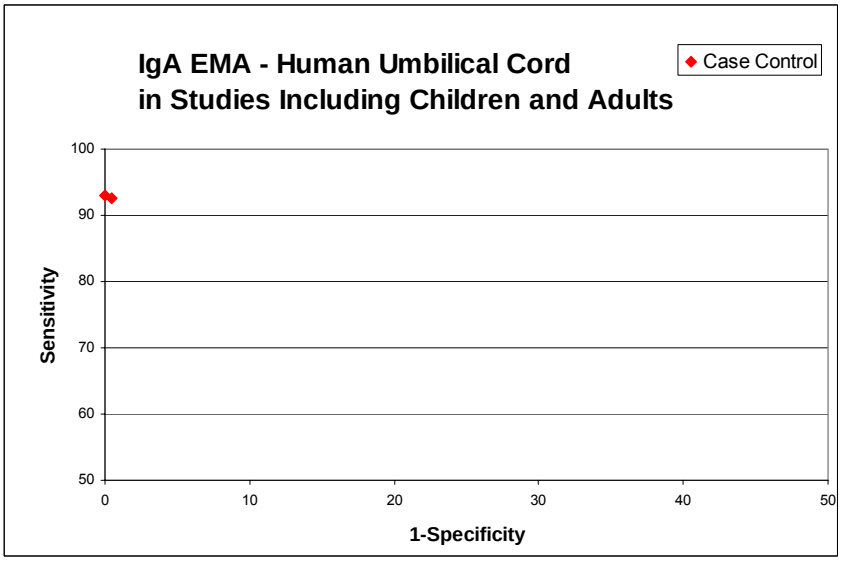


Figure 14: IgA-tTG-GP in adults with CD

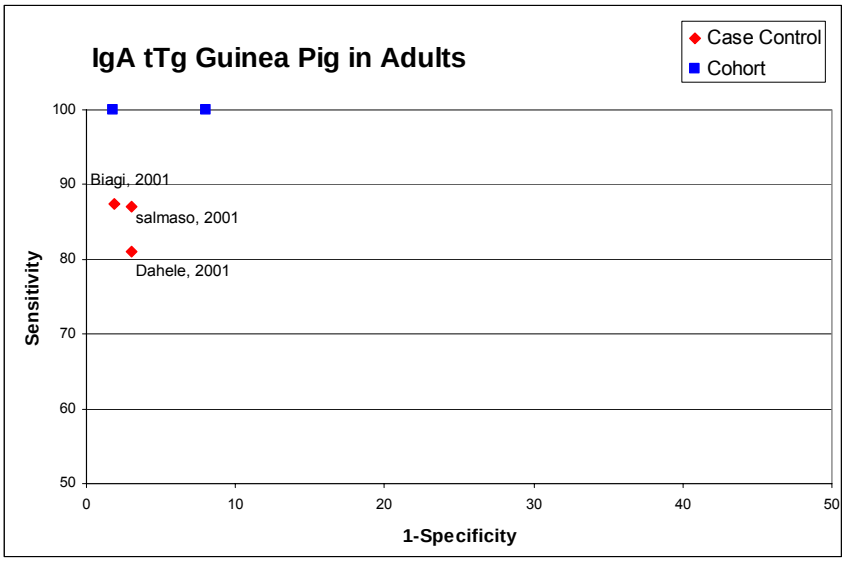


Figure 15: IgA-tTG-GP in children with CD

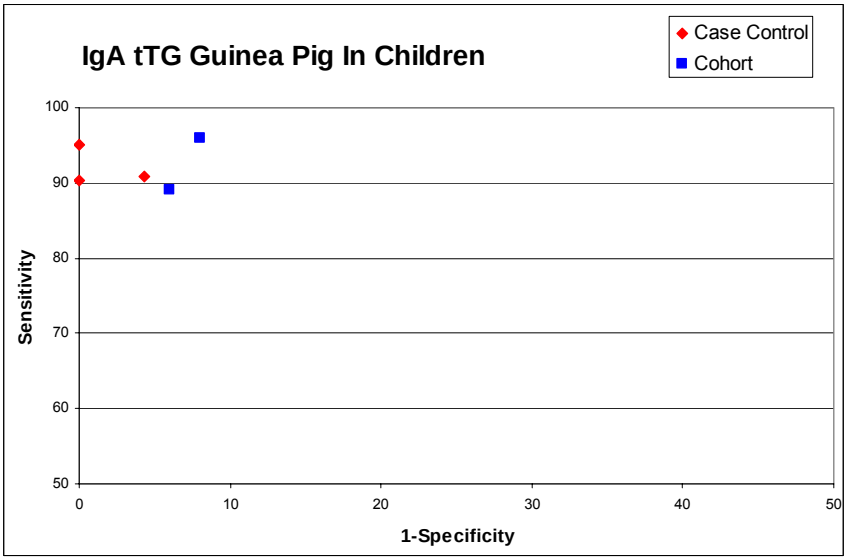
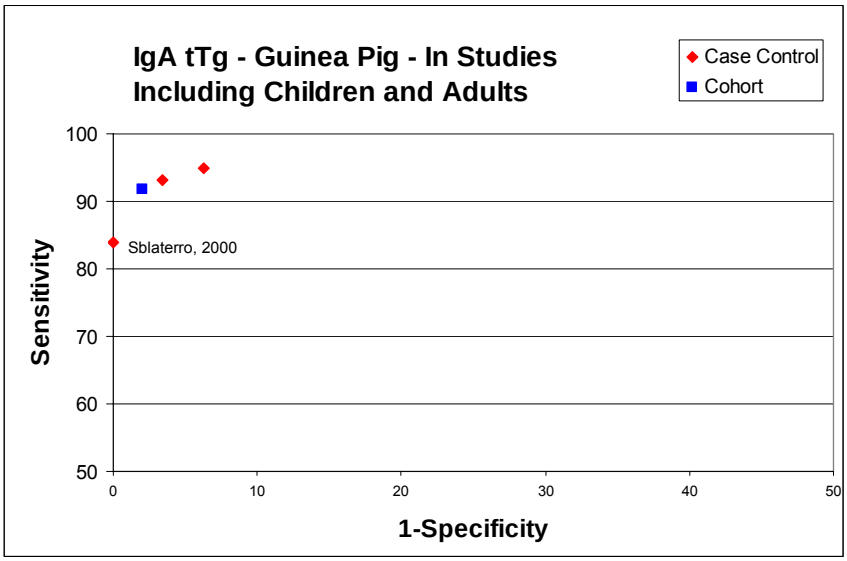


Figure 16: IgA-tTG-GP in adults and children with CD

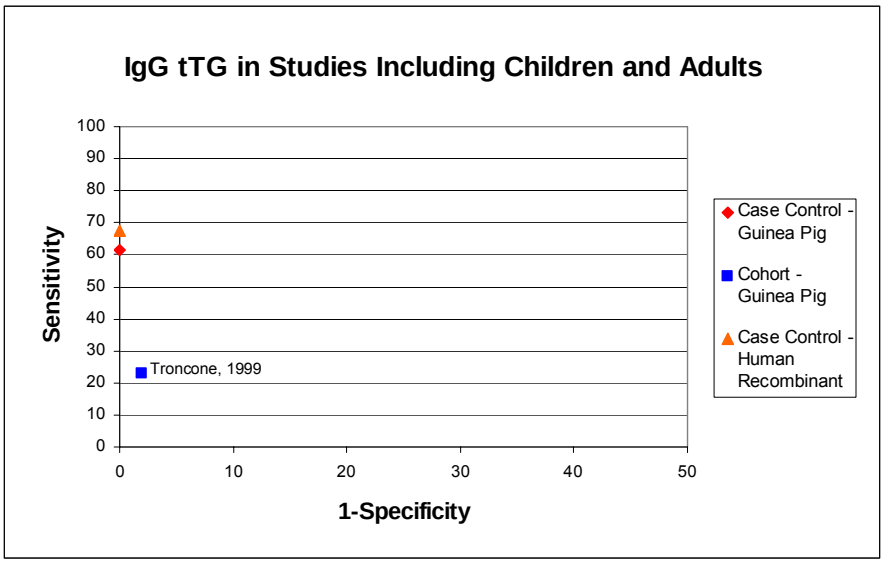


IgG-tTG-GP. Two studies in a mixed-age population assessed IgG-tTG- GP (Table 19; Figure 17).[72,76] The specificities in both studies were greater than 98%, but the sensitivities were 23% and 62%, respectively.

Table 19: Included studies for IgG-tTG-GP in studies including both children and adults

Author, year; country	Study type	Biopsy criteria	Sens	Spec	PPV	NPV	Prev
Sblaterro, 2000; Italy	Case-control	ESPGAN	61.5	100	100	44.4	0.8
Troncone, 1999; Italy	Relevant clinical population	ESPGAN	23	98	92	63	0.4

Figure 17: IgG-tTG-GP in adults with CD



tTG – human recombinant (HR)

IgG-tTG-HR. The diagnostic characteristics of IgA-tTG-HR were assessed by ELISA in ten studies, and the diagnostic characteristics IgG-tTG-HR were assessed by ELISA in two studies. Of the IgA-tTG-HR studies, three were conducted in adults,[39,49,54] three in children,[52,79,83] and three in a mixed population.[40,72,75]

Of the IgG-tTG-HR studies, two were conducted in a mixed population (Table 20),[40,72] but none were conducted in adults or children. One study was conducted in IgA-deficient patients and is described below.[40]

Two studies included CD patients with less than a Marsh IIIa grade.[45,70] These studies demonstrated sensitivities of 81% and 95% for IgA-tTG-GP.

One study was conducted in a mixed-age population of patients with known IgA deficiency.[40] In this study, the sensitivity of IgA-tTG–HR was 0%, wheras, the sensitivities and specificities of IgG-tTG-HR were 100% and 80%, respectively.

Table 23: Included studies for IgA-tTG-HR in studies including both children and adults

Author, year; country	Study type	Biopsy criteria	Notes	Sens	Spec	PPV	NPV	Prev
Cataldo, 2000; Italy	Case-control	Original & revised criteria?	20 IgA deficient CD vs healthy IgA-deficient non-CD	0	100	0	33.3	0.7
Sblaterro, 2000; Italy	Case-control	ESPGAN		91.5	100	100	76.9	0.8
Tesei, 2003; Argentina	Relevant clinical population	Marsh II to IV - with confirmation		91	96	97	87	0.6

Figure 18: IgA-tTG-HR in adults with CD

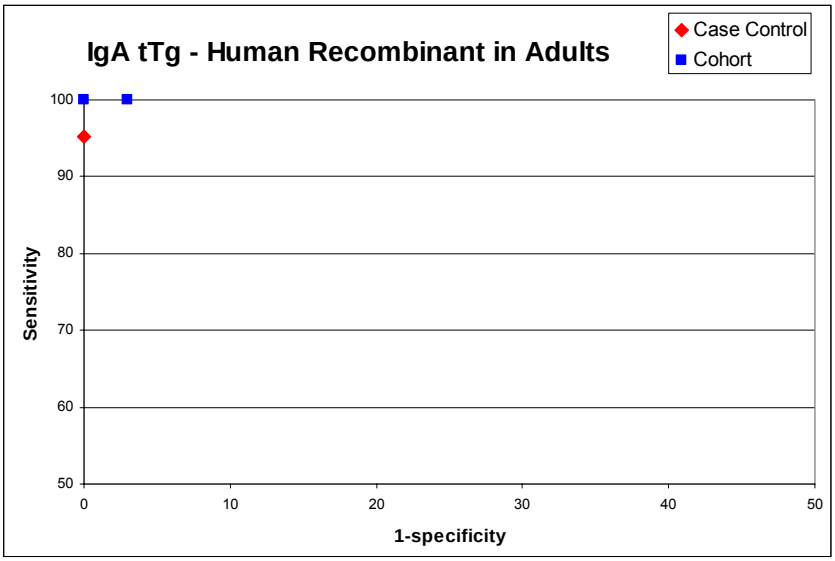


Figure 19: IgA-tTG-HR in children with CD

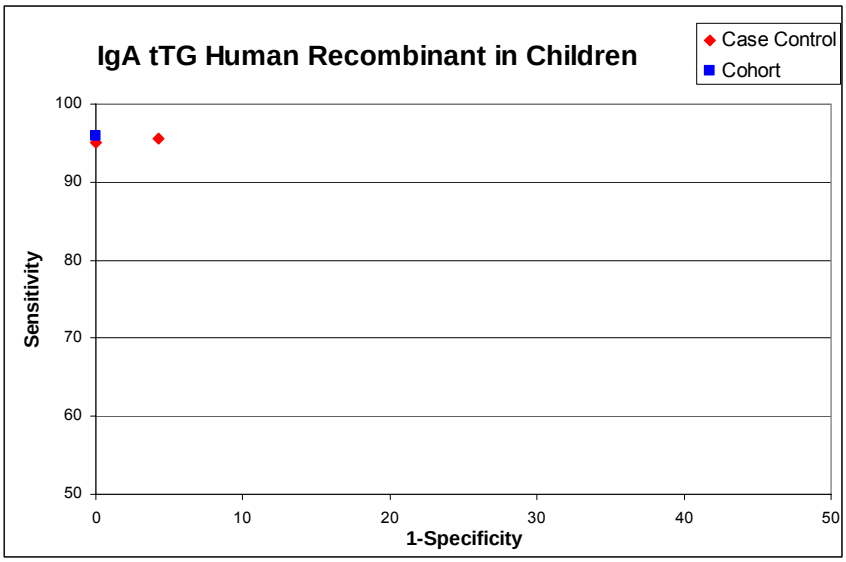
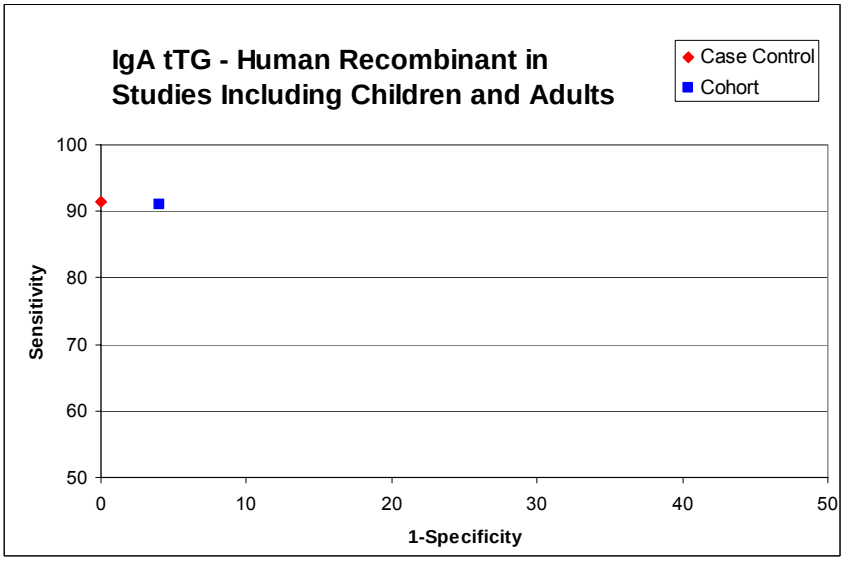


Figure 20: IgA-tTG-HR in adults and children with CD



IgG-tTG-HR, IgA deficient. Only one study of IgG-tTG-HR, conducted in an IgA-deficient population, was identified.[72] In this study, the sensitivity and specificity of IgG-tTG-HR was 68% and 100%, respectively.

Figure 21: PPV and prevalence from individual studies

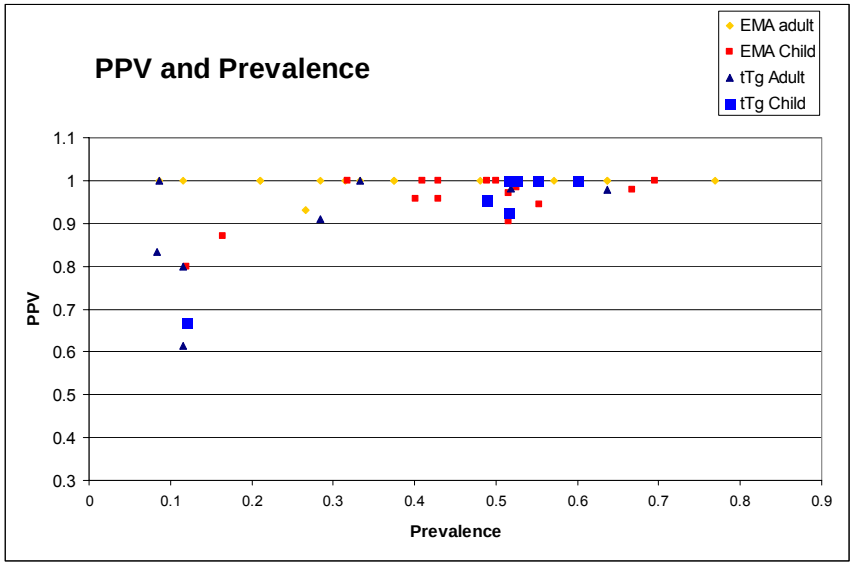


Figure 22: NPV and prevalence from individual studies

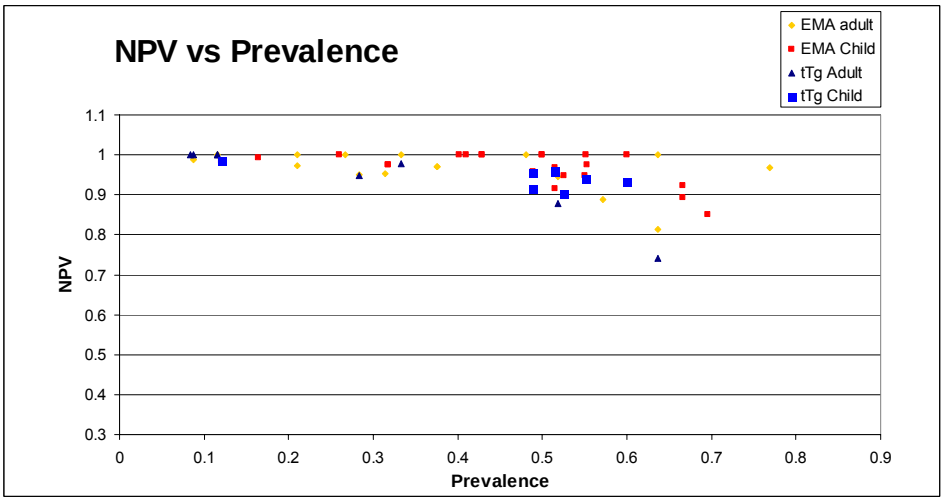


Figure 23: PPV based on the pooled estimates of sensitivity and specificity

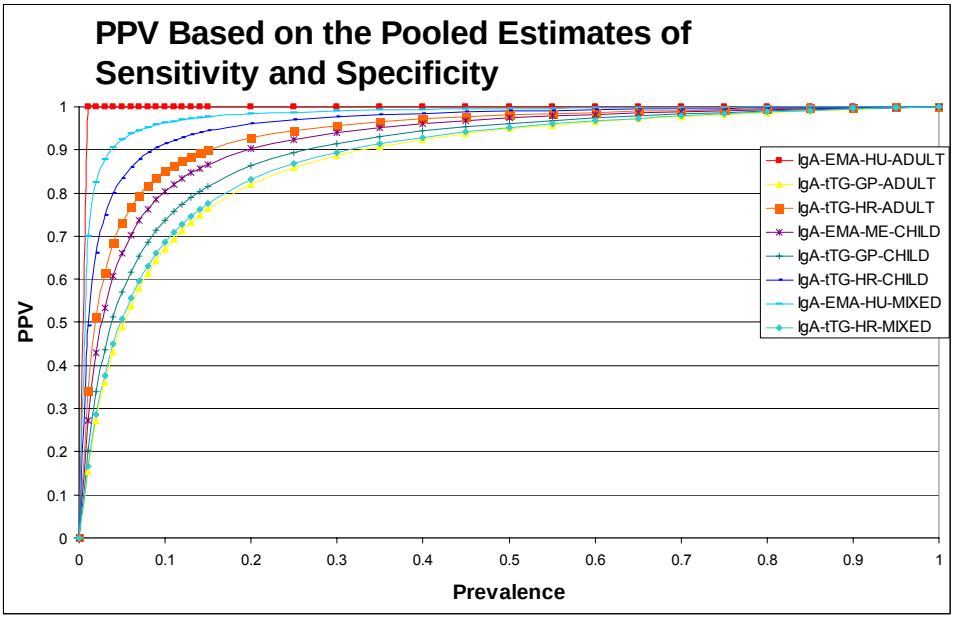


Table 33: Sensitivity/specificity (calculated) for HLA DQ2 or DQ8

Author; year; country	Prev of CD	DQ2 or DQ8 in CD	DQ2 or DQ8 in controls	Sens	Spec	PPV	NPV	Notes
Fasano, 2003; USA	0.52	100 (98/98)	59.78 (55/92)	1	0.4	0.64	1	Screened at-risk and not-at-risk populations
Catassi, 2001; Algeria	0.37	96.2 (76/79)	41.9 (57/136)	0.96	0.58	0.57	0.96	Saharawi Arabs
Lui, 2002; Finland	0.52	99.62 (259/260)	67.93 (161/237)	1	0.32	0.62	0.99	Family members of CD (controls=unaffected family members)
Balas, 1997; Spain	0.22	99.06 (210/212)	46.09 (342/742)	0.99	0.54	0.38	1	Known CD
Sumnik, 2000; Czech	0.07	100 (15/15)	87.63 (163/186)	1	0.12	0.08	1	Diabetes
Tuysuz, 2001; Turkey	0.52	90.91 (50/55)	32 (16/50)	0.91	0.68	0.76	0.87	Turkish Known CD
Neuhausen, 2002; Israel	0.31	100 (23/23)	86.54 (45/52)	1	0.13	0.34	1	Bedouin Arabs
Perez-Bravo, 1999; Chile	0.33	37.1 (23/62)	15.32 (19/124)	0.37	0.85	0.55	0.73	Chileans

Figure 24: HLA DQ2

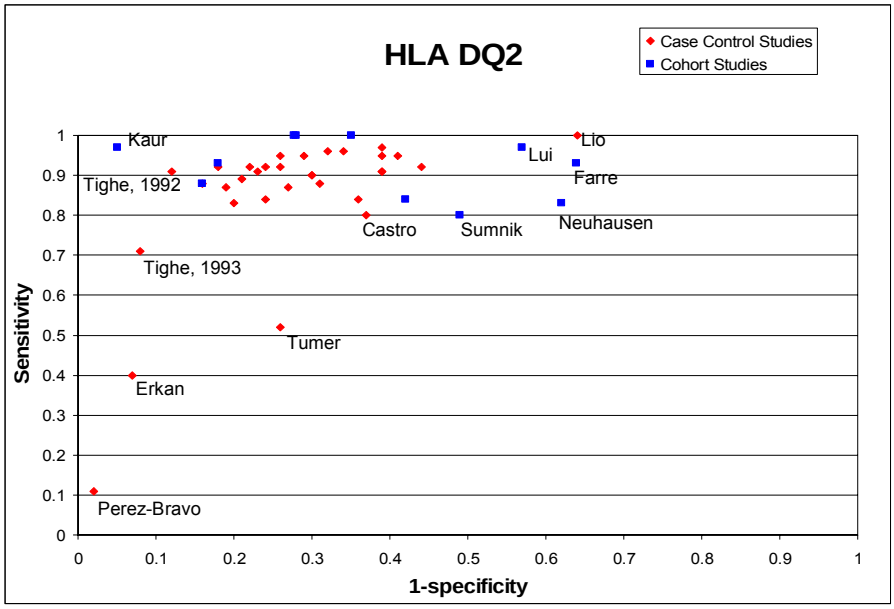
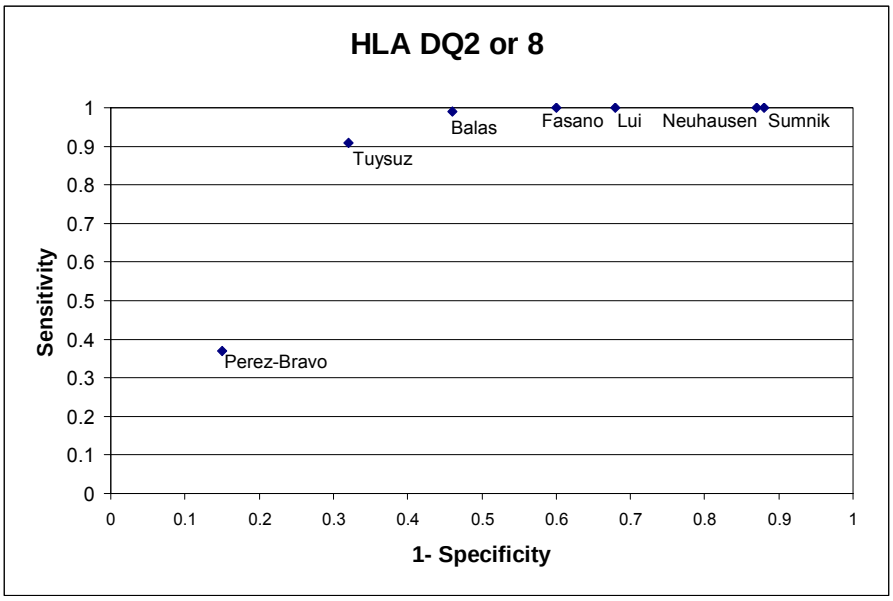


Figure 25: HLA DQ2 and DQ8



Biopsy

Using epidemiologically appropriate eligibility criteria, our comprehensive literature search did not identify any studies that specifically addressed the question of the sensitivity or specificity of biopsy for the diagnosis of CD.

However we sought to obtain indirect evidence regarding the diagnostic performance of biopsy as a test for CD. Some data was available from those studies identified for other review objectives, such as the cross-sectional screening studies, the HLA DQ2/8 studies, and studies of IELs. We also sought studies of follow-up of biopsy negative patients suspected of CD, and studies of silent and latent CD. The findings from these studies are presented in the Discussion and in Appendix H.

Quality Assessment

Overall, the quality of the diagnostic studies assessed in the Celiac 1 objective was quite good (Appendix J, Table 1). However, 59% of the studies reported using a selected patient population that may not be representative of a clinically relevant population. This is likely related to study design. Only 11% of the studies reported on whether the reference test was reported without knowledge of the index test. We felt that this was not a major threat to the validity of the studies.

Table 37: Prevalence of CD by statistical percentiles

Percentiles	Serology	Biopsy
5	.0016255	.0007378
10	.0018050	.0015761
25	.0030919	.0025321
50	.0063702	.0047672
60	.0084439	.0050768
75	.0117290	.0071429
80	.0125193	.0074416
90	.0184088	.0147536
95	.0225417	.0183992
100	.0266667	.0186722
Minimum	.00152	.00065
Maximum	.02667	.01867

Prevalence expressed as proportion (multiply by 100 for percent, or 100,000 for per 100,000 value)

Figure 26: Frequency distribution of prevalence of CD by serology among included studies

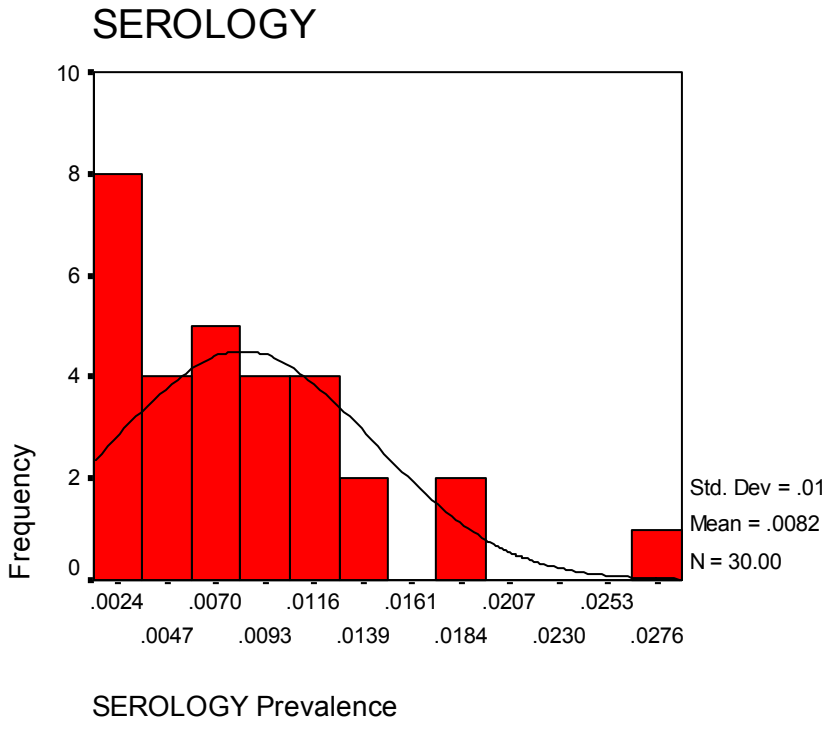


Figure 27: Frequency distribution of prevalence of CD by biopsy among included studies

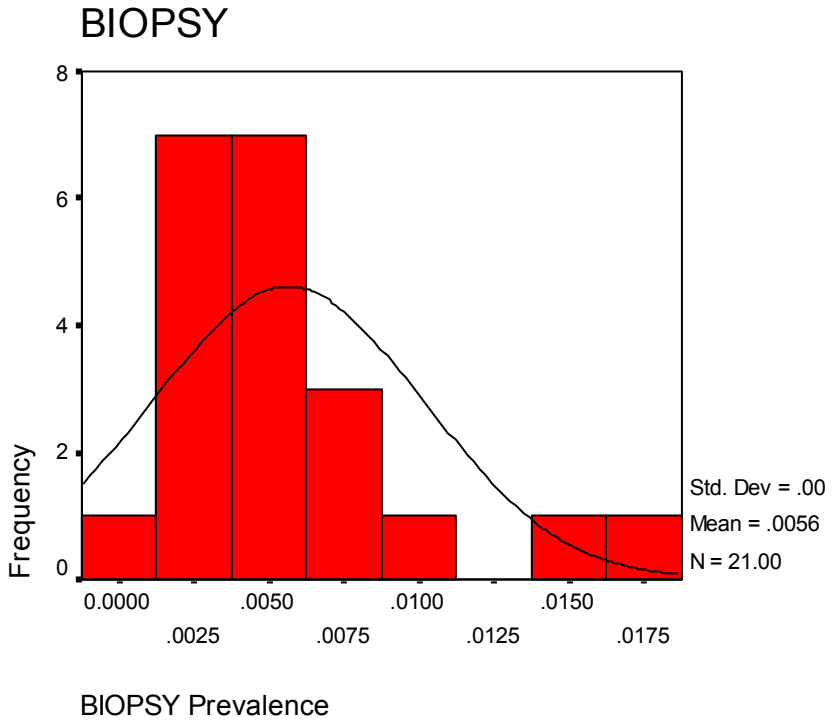


Figure 28: Prevalence of CD by country

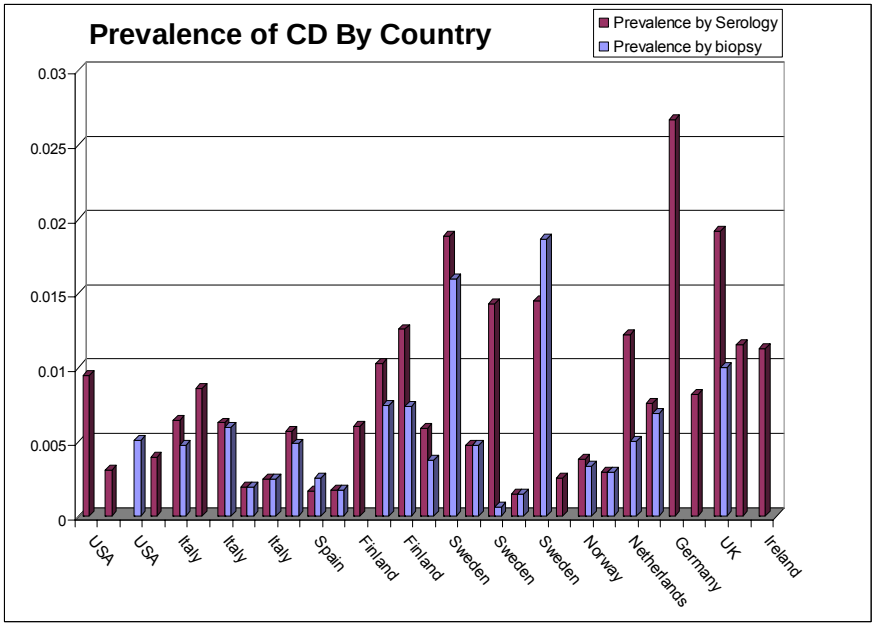
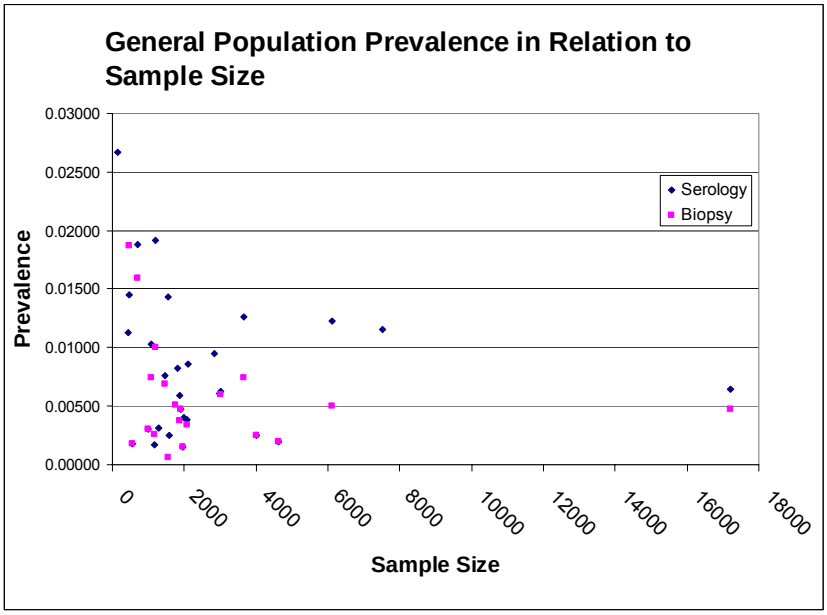


Figure 29: General population prevalence in relation to sample size



Prevalence of CD in Patients with Suspected CD

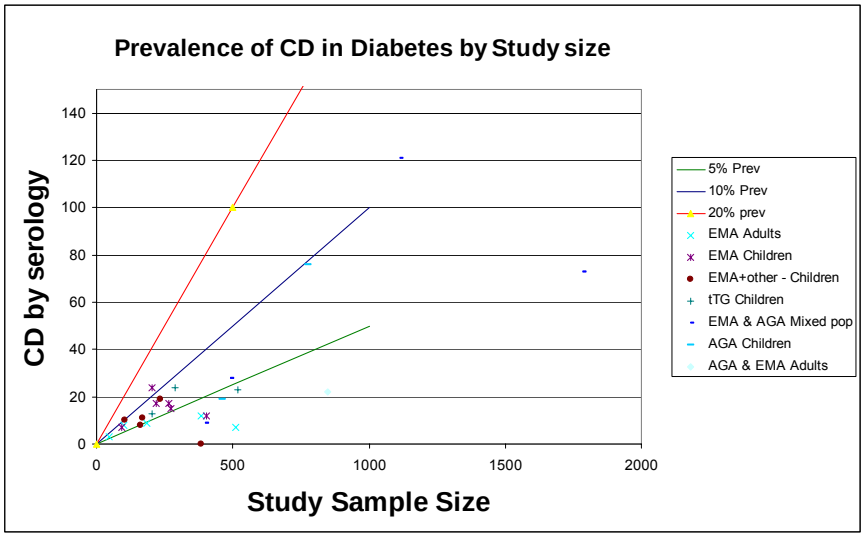
Adults: The prevalence of CD in adults suspected of the diagnosis was reported in four studies (Evidence Table 5, Appendix I; Table 38); three from Italy,[300,301,303] and one from the US.[206] The following reasons for suspecting a diagnosis of CD were documented: anemia, persistent iron deficiency, bowel disturbances, chronic intermittent diarrhea, abdominal pain, constipation, dyspepsia, severe malabsorption, tiredness and weight loss, mineral metabolism deficiencies, osteoporosis, arthralgias, arthritis, dermatitis, hypertransaminasemia, type I diabetes mellitus, infertility, and gluten intolerance in childhood not further investigated.

All three Italian studies were from referral centers, and intestinal biopsies were performed on all suspected cases, which cumulated to 347. The prevalence of CD was very high in these series, i.e., 43%,[300] 50%,[301], and 12%.[303]

In a large study of prevalence of CD in at-risk and not-at-risk individuals in the US, a total of 1,910 adults with CD-associated symptoms or disorders underwent serological testing with EMA. Fifteen of the 28 EMA-positive subjects (53.6%) consented to a biopsy, which was confirmatory in all cases.[206] The source of these patients and their mode of recruitment/referral were not reported. Based on the EMA result, the prevalence of CD in these adults with suspected CD was 1.5%.

Children: The prevalence of CD in children suspected of the diagnosis was reported in nine studies (Table 38); three from Canada,[304,305,307] two from the US,[206,324] and one each from Denmark,[302] England,[308] Italy,[309] and New Zealand.[306] The following reasons for suspecting a diagnosis of CD were documented: abdominal pain,[238,304,305,307,309] diarrhea,[238,304,305,308] failure to thrive/short stature,[206,238,304-306,309] weight loss,[305] vomiting,[304,305] abdominal distension,[304,305]

Figure 30: Prevalence of CD in diabetes by study size



Prevalence of CD in Relatives of Patients with CD

There were 18 studies on the risk of CD in first-degree relatives of patients with biopsy-proven CD,[129,167,206,235-249], four of which also provided data on the risk of CD in second-degree relatives (Evidence Table 7, Appendix I; Table 41).[206,235,238,239]

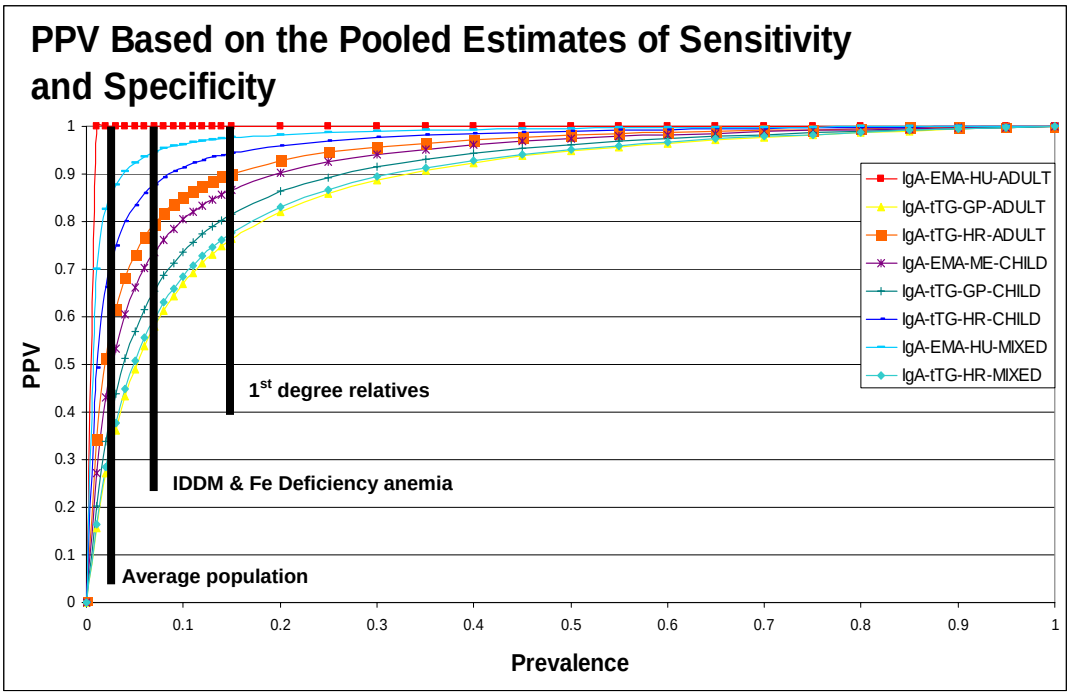
First-degree relatives: First-degree relatives were directly evaluated with small bowel biopsy in five studies; three were performed in England in the 1970's,[242,243,245] and two in Finland during the 1990's.[129,167] The biopsy criteria for a diagnosis of CD was not reported in one study,[243] and implied at least some degree of villous atrophy in the other four.[129,167,242,325] The percent of all at-risk family members that were studied varied from 34%[245] to 100%.[243] The study size varied between 29[242] and 182,[245] and the cumulative number of patients tested was 494. The prevalence of CD among first-degree relatives undergoing intestinal biopsy varied from 5.5%[243] to 22.5%;[245] the pooled prevalence was 16%.

Serological screening of the first-degree relatives of patients with biopsy-proven CD was performed in 12 studies.[206,235-237,239-241,244,246-249] In seven of those studies, intestinal biopsy was performed on at least 80% of the subjects who tested positive serologically, i.e., in 84 % of subjects in one study,[237] and in 100% of subjects in the other six studies.[236,239,244,247-249] Serological screening was performed with AGA alone in one study,[236] whereas, the other six studies used EMA, either alone[239] or in combination.[237,244,247-249] Six studies used criteria implying some degree of villous atrophy,[236,237,239,244,247,248] whereas, one study included cases with Marsh I changes.[249] The study size varied from 92[248] to 943[239] subjects, for a cumulative number of 2,607 subjects. For the studies that required some degree of villous atrophy for diagnosis, the prevalence varied from 4%[236] to 12%,[248] and the mean prevalence was 7.6%.

Information on the consequences of screening will come from the currently ongoing large population based prevalence studies.

The consequences of such issues as false-positive results were dealt with in the Celiac 1 Discussion. As discussed in that section, the definition of CD used and the prevalence of CD in the test populations, have a great impact on the diagnostic parameters of the available tests. We have presented data that show that the sensitivity of the available tests declines considerably when applied to patients with low-grade histological lesions. Unfortunately, there is insufficient data to address the question of what is the consequence of missing patients with low-grade histological lesions if serological screening alone is used. As described in Celiac 1, all the diagnositic test studies of the various serological markers were undertaken in study populations in which the prevalence of CD exceeded the that observed in most clinical situations. We have shown that the positive predictive value, which is predominately influenced by the test specificity and the prevalence of CD in the test population, drops from the reported values to much lower values when the test is applied in typical clinical populations. To illustrate this point, Figure 31 highlights the expected PPV when applied to different test populations.

Figure 31: PPV based on pooled estimates of sensitivity and specificity



As can be seen from Figure 31, the PPV—the probability that a positive test result actually represents true CD—drops with the prevelance of the population in which the test is applied. This relationship holds true for all the summary curves, but differ in degree. It is important to note that the PPV is predominantly influenced by the specificity of the test and prevalence. Since we have identified that the specificity of EMA and tTG is quite high, the major influence on the PPV in these analyses is the prevelance of CD in the population being tested. The practical importance of this discussion, is that despite having very high specificity, the use of these serological markers in low-prevalence populations would be expected to result in high false-positive rates. Below a prevalence of 5%, the false-positive rates may be as high as 30% to