Concordance, disease progression, and heritability of coeliac disease in Italian twins

Correspondence to:
Lorenza Nisticò and Corrado Fagnani
Genetic Epidemiology Unit
C-NESPS
Istituto Superiore di Sanità
Viale Regina Elena 299
00161 Roma, Italia
lnistico@iss.it
fagnani@iss.it

Lorenza Nisticò, Corrado Fagnani, Iolanda Coto, Selvaggia Percopo, Rodolfo Cotichini, Maria Giovanna Limongelli, Franco Paparo, Sandra D’Alfonso, Mara Giordano, Concettina Sferlazzas, Giuseppe Magazzù, Patricia Momigliano-Richiardi, Luigi Greco, Maria Antonietta Stazi.

L Nisticò, C Fagnani, R Cotichini, M A Stazi: Genetic Epidemiology Unit, National Centre of Epidemiology, Surveillance and Health Promotion (C-NESPS), Istituto Superiore di Sanità, Rome, Italy
I Coto, S Percopo, M G Limongelli, F Paparo, L Greco: Department of Paediatrics, University of Naples Federico II, Naples and European Laboratory of Food Induced Disease (ELFID), Naples, Italy.
S D’Alfonso, M Giordano, P Momigliano-Richiardi: Department of Medical Sciences, Eastern Piedmont University “A. Avogadro”, Novara, Italy
C Sferlazzas, G Magazzù: Department of Paediatric Sciences, University of Messina, Messina, Italy.

Key words: twin, coeliac disease, concordance, disease progression, heritability

Abbreviations: CD, coeliac disease; MZ, monozygotic, DZ, dizygotic; AIC, Italian Coeliac Disease Association; EMA, anti-endomysial antibodies, anti-tTG, anti-human-tissue-transglutaminase antibodies; ELISA, enzyme linked immunosorbent assay; ACE, additive genetic, common and unshared environmental factors; 95% CI, 95% Confidence Interval.
ABSTRACT

Background and aims: We adopted the twin method to disentangle the genetic and environmental components of susceptibility to coeliac disease (CD). We estimated disease concordance rate by zygosity and HLA genotypes, discordance times, progression rates to the disease and heritability.

Methods: We cross-linked the Italian Twin Registry with the membership lists of the Italian Coeliac Disease Association and recruited 23 monozygotic (MZ) and 50 dizygotic (DZ) twin pairs with at least one affected member. Zygosity was assigned by DNA fingerprinting and HLA-DQ and DR alleles were genotyped. Disease status was ascertained by anti-endomysial (EMA), anti-human-tissue-transglutaminase (anti-tTG) antibodies and bowel biopsy.

Results: Concordances were significantly higher in MZ (83.3% probandwise, 71.4% pairwise) than in DZ pairs (16.7% probandwise, 9.1% pairwise). Concordance was not affected by gender or HLA genotype of the co-twin and being MZ was significantly associated with occurrence of CD (Cox adjusted hazard ratio = 14.3; 95% CI: 4.0-50.3). In 90% of concordant pairs the discordance time was ≤ 2 years. MZ and DZ co-twins had respectively 70% and 9% cumulative probability to have symptomatic or silent forms of CD within 5 years. Under ACE models with CD population prevalences of 1/91 and 1/1000, the heritability estimates were 87% and 57%, respectively.

Conclusion: MZ pairs have high probability to be concordant regardless of gender or HLA genotype. Most of the affected co-twins receive the diagnosis within 2 years. A remarkable proportion of phenotypic variance is due to genetic factors.
INTRODUCTION

Coeliac disease (CD) is an immune-mediated intolerance triggered by the ingestion of gluten-containing grains in susceptible individuals.

It occurs with an inflammation of the upper small intestine, which may lead to the malabsorption of important nutrients, such as iron, folic acid, and calcium. Several symptoms may be related to untreated CD. The clinical presentation of these symptoms varies according to age and a great number of patients present the condition in a silent form; as a consequence, the population prevalence of CD is probably underestimated. Indeed, a large scale screening in the Italian paediatric population revealed the high prevalence of CD (1:91) with two third of the cases being clinically silent.[1]

From the etiological standpoint, CD is a multifactorial disease, involving both genetic and environmental factors. Besides the HLA region, already known to bear one or more risk loci for CD, at least two additional susceptibility loci map on chromosome 5q31-33 and 2q33.[2][3]

Despite a number of assumptions that have extensively been investigated in the past and might not be met in specific situations, the twin method constitutes a powerful tool to disentangle the genetic and environmental causes of family resemblance for a given trait.[4] Over the last years, the potential of twin studies has enormously increased with the establishment of population based registries of data on twins that represent some of the best resources for genetic epidemiologic research.[5] Matching one of such registries with disease records makes it feasible to collect twin pairs in which at least one member is affected; this results in samples that are largely representative of the general twin population, and also ensures a gain in terms of statistical power.

We adopted the twin approach to estimate: (1) CD concordance rate by zygosity and HLA status, (2) discordance times, (3) progression rates to the disease, and (4) heritability.

SUBJECTS and METHODS

Twin recruitment

Twins were identified according to the procedure already described in a previous paper.[6] Briefly, membership lists (6998 records of the Italian Coeliac Disease Association, AIC) were matched with the Italian Twin Registry that includes approximately 650,000 potential twin pairs born before 31 December 1996.[7][8] AIC lists were from Southern (Campania, Basilicata, Puglia, Calabria, Sicilia) and Northern (Piemonte) Italy regions. Eighty-one twin pairs were identified giving a ratio of 2.3 twins/100 individuals: 58 pairs were already known,[6] and 23 were identified de novo (13 from Piemonte and 10 from Southern Italy). Eight pairs (4 opposite sex, 3 female and 1 male) refused to participate and seventy-three pairs entered our study. In this paper the twin in a pair who first received CD diagnosis is referred to as the proband or the index twin; co-twin is sometimes indicated as the second twin. The study was approved by the Ethical Committee of the University of Naples and by the AIC. All twin pairs or their parents gave their informed consent.

HLA typing

Peripheral blood was drawn from 73 twin pairs and 128 parents, using EDTA as anticoagulant. Genomic DNA was purified from leukocytes of all subjects with the salting out procedure.[9] HLA-DRB1 and -DQB1 loci were typed in all twins and their parents, when available, using commercial kits (Dynal Oxoid, Italy). HLA chromosomes identical by descent could be unambiguously assigned in 41 pairs.
Zygosity test

Zygosity was provisionally assigned by a standardised questionnaire and then confirmed in same sex twin pairs with identical HLA genotypes by DNA fingerprinting. We used the AmpFLSTR Profiler Plus PCR Amplification Kit (Applied Biosystems) that amplifies, at the same time, 9 Short Tandem Repeats all localised on different chromosomes. Probability of identity by chance is approximately $10^{-4}$. PCR products were separated by capillary electrophoresis on an automatic sequencer ABI Prism 310 (Applied Biosystems).

Disease status

Seemingly unaffected co-twins were visited and screened for CD: serum samples were examined for anti-endomysial (EMA) and anti-human-tissue-transglutaminase (anti-tTG) antibodies. IgA endomysial antibodies were detected by an indirect immunofluorescence method using cryostat sections of human umbilical cord as antigen;[10] sera were tested at 1:5 dilution. Serum IgA anti-tTG levels were measured by an ELISA method.[11] All samples were tested in duplicate and the values were expressed as percentage of a reference pool of sera, obtained from untreated coeliac patients. All unaffected co-twins were tested for IgA levels to exclude IgA deficiency. In all cases of positive screening, diagnosis of CD was confirmed by a small bowel biopsy showing typical pathological changes. All affected subjects fulfilled the ESPGHAN (European society of paediatric gastroenterology, hepatology and nutrition) diagnostic criteria.

Statistical Analysis

1. Concordance

We estimated CD concordance by zygosity and sex, using probandwise ($P_C$) and pairwise ($P_P$) concordance rates under incomplete ascertainment.[12] Concordant affected pairs were distinguished between “doubly” (D) ascertained, for which both twins were in the disease records, and “singly” (S) ascertained, where only one twin was in the records, and the second twin was found to be affected on further examination. $P_C$ is defined as the probability that one twin in a pair is affected, given that his/her co-twin is affected, and can be estimated as: $P_C = (2D+S)/(2D+S+d)$, where $d$ is the number of discordant pairs. $P_P$ is the probability that both twins in a pair are affected, given that at least one is affected, and can be expressed as: $P_P = (2D+S)/(2D+S+2d)$.

2. Survival

We performed survival analyses by the Kaplan-Meier method on 72 pairs, to describe the progression of CD in co-twins. One MZ concordant pair was excluded because both twins were screened shortly after CD was diagnosed in another sibling. We considered two survival models: one in which the terminating event was defined as CD diagnosis as consequence of appearance of symptoms or positive screening and biopsy in silent co-twins, and another one where the event was defined as CD diagnosis as consequence of symptoms appearance only, while positively screened twins were incorporated as censored observations. Moreover, we adopted the Cox regression model to evaluate the impact of co-twin sex, HLA genotype group and zygosity on CD concordance.

3. Heritability

We estimated genetic and environmental components of variance in CD using structural equation modeling and the software Mx.[13] We considered an ACE model incorporating parameters for additive genetic (A), common (shared) environmental (C), and individual-specific (unshared)
environmental (E) components of variance.[4] Additive genetic factors are completely correlated in MZ twins, who are genetically identical, and correlate 0.5 in DZ twins, who share on average 50% of their segregating genes. Common environmental factors are shared completely by the twins regardless of zygosity, while unshared environmental influences act separately on each twin and therefore are responsible for less than perfect resemblance between MZ twins. Under these assumptions, the expectations for the total variance (V) and the covariance within twin pairs are given by:[4]

\[
V = A + C + E \\
Cov(MZ) = A + C \\
Cov(DZ) = 0.5*A + C
\]

Heritability (h²) is the proportion of the total variance that is attributable to the genetic variance; in formula: \( h^2 = A/V \).

The A, C and E parameters were estimated by maximum likelihood, under a “liability-threshold” model:[14] based on this model, there exists an underlying (continuous) liability to CD, normally distributed in the population, with a threshold such that individuals above the threshold are affected and those below are not.

This study relied on incomplete ascertainment of twins.[12] Thus, the likelihood estimation had to be corrected by taking into account the ascertainment probability (0<\( \pi <1 \)) (i.e. the probability that an affected twin is in the disease records) and by fixing the threshold of liability at the value corresponding to the population prevalence of the disease. The ascertainment probability can be approximated as \( \pi = 2D/(2D+S) \) where D and S are numbers of concordant pairs doubly and singly ascertained as described in the Concordance section. The appropriate script is available on the Mx website: http://www.vcu.edu/mx/examples.html.

**RESULTS**

**Recruitment and zygosity**

We enrolled 73 twin pairs. Twenty-three pairs were MZ (6 males, 17 females) and 50 were DZ (12 male, 15 female and 23 opposite sex). The MZ/DZ same sex/DZ opposite sex ratio was 0.9/1.1/0.9 and was not significantly different from that expected (1/1/1). Female/male ratio of the index twins was 2.2. Age at enrolment was similar in MZ and DZ twins (Table 1; t-test: p = 0.30).

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Ages at diagnosis and at enrolment of the twins</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Monozygotic twins</td>
</tr>
<tr>
<td>No of pairs</td>
<td>Concordant</td>
</tr>
<tr>
<td>Mean age at diagnosis (y) in the index twin</td>
<td>10.3</td>
</tr>
<tr>
<td>Mean age at diagnosis (y) in the co-twin</td>
<td>11.3</td>
</tr>
<tr>
<td>Mean age at enrolment (y)</td>
<td>21</td>
</tr>
</tbody>
</table>

**Disease status**

Seventeen pairs were known to be concordant before entering the study. Five co-twins (3 female MZ, 1 male MZ and 1 male DZ) were diagnosed during the study because they were positive to autoantibody screening and to intestinal biopsy. Four of them were clinically silent and one MZ co-twin had symptoms. Age at diagnosis varied greatly (range 0.5-57 years), although 50% of all
affected twins were diagnosed within 3 years of age. Mean age at diagnosis was similar in MZ and DZ index and second twins (Table 1) but was lower in male than in female twins (7.3 versus 10.8 years, t-test: p = 0.17).

When patients or their parents were asked about symptoms that led to diagnosis, the most frequent answers were: diarrhoea (51%), vomiting (39%), weight loss (29%), anaemia (22%) and abdominal distension (19%). Twelve twins claimed to be symptom-less and to have been positively screened for CD because of affected index twins (10 co-twins), affected mother or non-twin sister (2 index twins).

Concordance estimates

Overall, 17/23 MZ and 5/50 DZ twin pairs were concordant for CD: probandwise (Pc) and pairwise (Pp) concordances were significantly different between MZ and DZ twins, with point estimates of 83% and 71% in MZ twins, and 17% and 9% in DZ twins, respectively (Table 2). Concordances by gender did not significantly differ in MZ twins. In DZ, 1/5 concordant pairs was female and 4/5 were opposite sex. None of the 12 DZ male pairs was concordant (Table 2). In 15/19 discordant opposite sex pairs the affected twins were females.

<table>
<thead>
<tr>
<th></th>
<th>Concordant</th>
<th>Discordant</th>
<th>Total</th>
<th>Probandwise</th>
<th>95% CI</th>
<th>Pairwise</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>MZ male</td>
<td>5</td>
<td>1</td>
<td>6</td>
<td>90</td>
<td>70.6-100</td>
<td>81.8</td>
<td>49.7-100</td>
</tr>
<tr>
<td>MZ female</td>
<td>12</td>
<td>5</td>
<td>17</td>
<td>80.8</td>
<td>64.4-97.1</td>
<td>67.7</td>
<td>44.7-90.7</td>
</tr>
<tr>
<td>All MZ</td>
<td>17</td>
<td>6</td>
<td>23</td>
<td>83.3*</td>
<td>70.3-96.4</td>
<td>71.4†</td>
<td>52.3-90.6</td>
</tr>
<tr>
<td>DZ male</td>
<td>0</td>
<td>12</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>DZ female</td>
<td>1</td>
<td>14</td>
<td>15</td>
<td>12.5</td>
<td>0-34.7</td>
<td>6.7</td>
<td>0-19.3</td>
</tr>
<tr>
<td>DZ opposite sex</td>
<td>4</td>
<td>19</td>
<td>23</td>
<td>26.9</td>
<td>5.2-48.7</td>
<td>15.6</td>
<td>1-30.1</td>
</tr>
<tr>
<td>All DZ</td>
<td>5</td>
<td>45</td>
<td>50</td>
<td>16.7*</td>
<td>3.6-29.8</td>
<td>9.1†</td>
<td>1.3-16.9</td>
</tr>
<tr>
<td>Total</td>
<td>22</td>
<td>51</td>
<td>73</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Test for difference between MZ and DZ twins: *χ² = 49.98 p = 1.55 x 10⁻¹²; †χ² = 40.77 p = 1.71 x 10⁻¹⁰

Tables 3 and 4 show disease concordance according to HLA genotypes in MZ and DZ twin pairs. It was recently shown that HLA-CD association is better described by a risk hierarchy of DR-DQ genotypes rather than by DQ2-DQ8 molecules.[15] Accordingly, we stratified twin pairs in four genotype groups with decreasing risk for CD: the highest risk is due to DR3/3 and DR3/7 genotypes (Group 1, G1); DR5/7 confers one-third less risk (G2); the relative risks of DR3/X heterozygous (G3) and of DR7/7, DR4/4 and DR4/7 genotypes (G4) are estimated to be approximately one-fourth of the G1 risk; finally, the fifth group (G5) includes genotypes (DR4/Y, DR5/X, DR7/Z) with very low risk of CD (1/50th of the G1).
Table 3 Concordance by HLA-DR genotype risk group in MZ pairs

<table>
<thead>
<tr>
<th>Risk genotype group†</th>
<th>Concordant</th>
<th>Discordant</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1 (DR3/3, DR3/7)</td>
<td>5</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>G2 (DR5/7)</td>
<td>3</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>G3-G4 (DR3/X, DR4/4, DR4/7, DR7/7)</td>
<td>8</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>G5 (DR4/Y, DR5/X, DR7/Z)</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>17</td>
<td>6</td>
<td>23</td>
</tr>
</tbody>
</table>

† HLA grouping is according to relative risks for coeliac disease estimated in the Italian population in Margaritte-Jeannin et al.[15] DR3 haplotype carries DQB1*0201 and DQA1*0501, DR7 carries DQB1*0202, DR5 carries DQA1*0501 and DR4 carries DQB1*0302.

Table 4 Concordance in DZ pairs by HLA-DR genotype risk group in index and co-twin

<table>
<thead>
<tr>
<th>Risk genotype group in the index twin†</th>
<th>G1</th>
<th>G2</th>
<th>G3-G4</th>
<th>G5</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1 (DR3/3, DR3/7)</td>
<td>1</td>
<td>5</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>G2 (DR5/7)</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>G3-G4 (DR3/X, DR4/4, DR4/7, DR7/7)</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>G5 (DR4/Y, DR5/X, DR7/Z)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>2</td>
<td>9</td>
<td>0</td>
<td>4</td>
</tr>
</tbody>
</table>

† HLA grouping is according to relative risks for coeliac disease estimated in the Italian population in Margaritte-Jeannin et al.[15] DR3 haplotype carries DQB1*0201 and DQA1*0501, DR7 carries DQB1*0202, DR5 carries DQA1*0501 and DR4 carries DQB1*0302.

‡ One co-twin is DR13/13.

All MZ twin pairs, but one, belonged to G1-G4 risk categories (Table 3). Discordant MZ co-twins were DR3/7 (N = 2), DR5/7 (N = 3) and DR3/5 (N = 1). In DZ pairs, 48/50 index twins carried G1-G4 genotypes (Table 4) and 2/50 were DR1/7 and DR4/13 (G5).

In MZ, the highest proportion of concordant pairs was observed in the G3-G4 genotype group (8 concordant, 1 discordant). In DZ, the same proportions of concordant pairs were observed in the groups with both twins having high risk G1 genotypes (1 concordant and 5 discordant) or G3-G4 genotypes (2 concordant, 10 discordant).

In each HLA stratum, disease concordance was higher in MZ pairs compared to DZ sibs sharing identical HLA chromosomes (data not shown). Indeed, in the Cox regression model, being MZ was the only factor significantly associated with the occurrence of CD in the co-twin (adjusted hazard ratio by sex and genotype group in the co-twin = 14.3; 95% CI: 4.0-50.3).

Survival analysis

There were 8 disease concordant pairs (7 MZ, 1 DZ) in which symptoms appeared almost simultaneously (0-2 months) in both twins, and one DZ pair with a discordance time of 7 months. In 2 MZ pairs, diagnosed before the introduction of autoantibody screening, symptom appearance in the co-twins was delayed 2 years. In 10 concordant pairs the second twins were clinically silent or paucisymptomatic. Of these, 7 co-twins (5 MZ, 2 DZ) were screened for autoantibodies and
diagnosed as coeliac within 1 year from the diagnosis in their siblings, one MZ was diagnosed within 2 years and in 2 co-twins (1 MZ and 1 DZ) affection status was ascertained after 10 and 37 years. Thus, in most (19/21) of our concordant pairs the disease discordance time was ≤ 2 years, either as a consequence of symptom appearance or because disease signs were actively searched in clinically silent co-twins.

The cumulative probability of being diagnosed with CD was significantly higher for MZ compared to DZ co-twins in both survival models (symptomatic CD, log-rank test: p = 1.6x10^{-5}; symptomatic or silent CD, log-rank test: p = 3.5x10^{-9}).

In MZ co-twins the probability to have symptomatic forms of CD within 1 year after the diagnosis in their identical siblings was 36%, much higher than that observed in DZ co-twins (4%). The 5-year cumulative probability further increased for MZ (50%) while did not change for DZ co-twins (figure 1a). Including silent forms, the progression rates for MZ and DZ co-twins were 50% and 8% within 1 year, 70% and 9% within 5 years (figure 1b). In our discordant pairs the shortest follow-up times were 4.6 years for MZ and 1.8 years for DZ co-twins.

**Heritability**

In the present study we had 13 and 4 doubly ascertained pairs among MZ and DZ twins, respectively, and 4 and 1 singly ascertained pairs among MZ and DZ twins, respectively; this resulted in an ascertainment probability of \( \pi = (2*17)/(2*17+5) = 34/39 = 0.87 \).

Given the so-called “iceberg” structure of CD,[16] which underlines the importance of the subclinical component, we fitted the same ACE model assuming for the population prevalence the values 1/1000 (threshold 3.09) from clinical diagnosis data, and 1/91 (threshold 2.29) from screening data.[1] Each threshold is the value such that the proportion of the standard normal distribution of liability above the value exactly matches the observed proportion of affected individuals in the population, namely CD prevalence.

The estimates of the genetic and environmental variance components are shown in Table 5.

<table>
<thead>
<tr>
<th>Variance Components</th>
<th>Prevalence (threshold)</th>
<th>A (95% CI)</th>
<th>C (95% CI)</th>
<th>E (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/1000 (3.09)</td>
<td>0.57 (0.32-0.93)</td>
<td>0.42 (0.06-0.67)</td>
<td>0.01 (0.00-0.03)</td>
<td></td>
</tr>
<tr>
<td>1/91 (2.29)</td>
<td>0.87 (0.49-1.00)</td>
<td>0.12 (0.00-0.49)</td>
<td>0.01 (0.00-0.05)</td>
<td></td>
</tr>
</tbody>
</table>

Under the ACE model with population prevalence for CD of 1/1000, 57% of the variation in liability to CD was explained by additive genetic factors, while 42% and 1% were the contributions of common and unique environmental factors, respectively. When we fitted the same ACE model with a population prevalence of 1/91, the heritability estimate became 87%, and the relative weight of common environmental factors decreased to 12%; the contribution of the unshared environmental component of variance remained at the 1% level.

**DISCUSSION**

As expected, concordances in MZ twins were significantly higher compared to DZ pairs and were very close to those previously reported on smaller sample size.[6] Moreover, given that number of singly ascertained pairs was relatively small, concordance estimates were similar under both
complete and incomplete ascertainment scenarios. In DZ, proportion of affected co-twins (5/50 = 10%) was in line with CD prevalences of 4% to 12% in first-degree relatives reported in several studies.[17]

Concordance rates by gender were not significantly different within MZ and DZ groups. In DZ pairs, the highest point estimates in opposite sex twins were not due to mean follow-up time since it was similar to that of female pairs and even shorter than that of male pairs.

Six out of 23 MZ pairs were disease discordant after a period of 4.5-27 years of follow-up. Discordance in MZ twins is usually attributed to differential exposure to environmental risk factors; in the case of CD, precautionary gluten elimination from the diet of unaffected co-twins cannot be ruled out. An alternative explanation for the differences between genetically identical individuals involves epigenetic modifications, such as DNA methylation and histone acetylation, occurring after twin separation, that may control expression and silencing of disease genes.[18]

In both MZ and DZ pairs, we observed that proportions of affected co-twins were not significantly different between those with G1 or G3-G4 HLA genotypes. This may indicate either random fluctuation or that risk hierarchy shown in the general population cannot be applied to this specific setting where high-susceptibility HLA genotypes may have been picked up.[15]

In our co-twins, the only significant susceptibility factor to CD is sharing the whole genetic background with the affected twin. Indeed, at least two additional loci on chromosomes 2 and 5 are associated with or linked to CD in the Italian population.[2][3]

In our concordant pairs, most of the co-twins received CD diagnosis shortly after their siblings, with median discordance time of 1 month for both MZ and DZ twins; since median follow-up times in discordant pairs were 10.5 and 7.7 years for MZ and DZ, respectively, we do not expect any dramatic variation of the concordance rates as a consequence of an even longer follow up.

It is well known that CD is frequently paucisymptomatic or silent. In our study, approximately half of MZ and DZ affected co-twins were positive to antibody screening and had flat gut mucosa despite the scarcity or absence of symptoms. Thus, if CD diagnoses due to symptoms appearance only had been considered, disease incidence would have been largely underestimated in MZ and DZ non-index twins. The 5-year cumulative incidence ratios of MZ relative to DZ co-twins were 12.5 (0.50/0.04) when excluding silent co-twins and 7.8 (0.70/0.09) if asymptomatic second twins were included.

Our data showed a substantial heritability for CD. Under ACE models with population prevalences for CD of 1/1000 and 1/91, the heritability estimates were 57% and 87%, respectively, indicating that approximately sixty to ninety percent of the variance in liability to the disease has a genetic origin. However, because of the limited power of this study, which reflects on large confidence intervals, we have to be cautious in interpreting the point estimates.

Moreover, it could be worth reminding the reader that these estimates give us only a general understanding of the genetic influence on the phenotype. That is, they simply suggest a genetic role in determining inter-individual phenotypic differences, but they tell us nothing about which genes are directly implicated in the emergence of the disease. And they give no information at all on whether it is possible to modify the prognosis by environmental interventions, such as lifestyle changes. With respect to this issue, the elimination of gluten from the diet is a typical example of environmental intervention that, in the case of CD, can result in a total recovery of gut function and a correction of most other consequences, despite a considerable heritability.

It is clear that the approach we used strongly depends on fixing a priori the level of the population prevalence; this is required to correct the likelihood function for the ascertainment procedure when selected samples of twins, including only pairs with at least one affected member, are considered. In
our case, if we had solely relied upon the prevalence estimated from clinical data, we would have underestimated the genetic penetrance and overestimated the shared environmental influences.

Based on our data, common environment could not completely be ruled out as a contributing etiological factor. Common environment refers to any shared environmental factor that contributes to the resemblance of members of a twin pair regardless of zygosity. It may include biological events like exposure to infectious agents, dietary characteristics, and other intra-uterine factors that may influence the similarity of CD patterns within pairs. Moreover, it is clear that the environments of members of twin pairs are most alike in utero and during the first post-natal period, while tend to diverge over time. Therefore, possible effects of common environment on liability to CD could be seen both in the short discordance time and in the young age at diagnosis observed, in our sample, for the majority of concordant pairs.

To our knowledge, no previous investigations have aimed at estimating, by the twin model, the heritability of CD in Italy or in other Countries. Rather, family studies have been extensively used for unravelling the genetic predisposition to CD.[19][20] In principle they consist of comparing the risk for relatives of affected individuals with that in the general population, and unlike the twin method they suffer the disadvantage of not providing information on to what extent the observed familial resemblance has a genetic basis or is attributable to shared environmental exposures.

The heritability point estimate under high prevalence scenario was quite in line with the figures found in twin studies on other HLA-mediated diseases, such as type 1 diabetes (88%),[21] Graves’ disease (79%) and psoriasis (80%);[22][23] this plausibly reflects shared pathogenetic mechanisms that may partially explain the co-morbidity of these conditions.[24][25][26]

ACKNOWLEDGEMENTS

We warmly thank the twins, their families and the Associazione Italiana Celiachia (AIC).

We are grateful to the following clinical investigators: Giuseppina Oderda (Dipartimento Scienze Mediche, Università del Piemonte Orientale), Cristiana Barbera (Clinica Pediatrica, Università di Torino), Carla Sategna-Guidetti (Dipartimento Medicina Interna, Università di Torino), Giuseppe Di Pasquale and Riccardo Scoglio (Dipartimento Scienze Pediatriche, Mediche e Chirurgiche, Università di Messina). We also acknowledge the skilful assistance of Cristina D’Ippolito in the management of the Italian Twin Registry database.

This work was supported by ELFID (European Laboratory for Food Induced Disease), the Italian Ministero della Salute (Project 2000, number 0AB/F to M.A.S.) and the Italian Ministero dell’Istruzione, dell’Università e della Ricerca (MIUR-PRIN no. MM06187812 to S.DA., G.M. and L.G.).

C.F. was supported by the Genoomeutwin project (European Union Contract No. QLG2-CT-2002-01254).

The Authors declare to have no competing interests.

Ethics approval: the study was approved by the Ethical Committee of the University of Naples and by the Italian Coeliac Disease Association.

The Corresponding Authors have the right to grant on behalf of all authors and does grant on behalf of all authors, a non exclusive licence on a worldwide basis to the BMJ Publishing Group Ltd and its Licensees to permit this article to be published in Gut editions and any other BMJPGL products to exploit all subsidiary rights, as set out in our licence.
REFERENCES


Legend to Figure 1. Kaplan-Meier survival curves in co-twins of coeliac probands by zygosity. The terminating event is: a) CD diagnosis as consequence of symptoms appearance only; b) CD diagnosis as consequence of either appearance of symptoms or positive screening. NDZ and NMZ are the numbers of co-twins entering each time interval.
### Cumulative Survival

#### Time of discordance (years)

<table>
<thead>
<tr>
<th></th>
<th>MZ</th>
<th>DZ</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDZ</td>
<td>50</td>
<td>32</td>
</tr>
<tr>
<td>NMZ</td>
<td>22</td>
<td>6</td>
</tr>
</tbody>
</table>

#### Time of discordance (years)

<table>
<thead>
<tr>
<th></th>
<th>MZ</th>
<th>DZ</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDZ</td>
<td>50</td>
<td>32</td>
</tr>
<tr>
<td>NMZ</td>
<td>22</td>
<td>6</td>
</tr>
</tbody>
</table>