Coeliac disease and risk of tuberculosis: a population based cohort study

J F Ludvigsson, J Wahlstrom, J Grunewald, A Ekbom, S M Montgomery

Background: Coeliac disease (CD) is an autoimmune disease often characterised by malnutrition and linked to a number of complications such as an increased risk of lymphoma, adverse pregnancy outcome, and other autoimmune diseases. Tuberculosis (TB) affects a large proportion of the world population and is more common in individuals with malnutrition. We investigated the risk of TB in 14 335 individuals with CD and 69 888 matched reference individuals in a general population based cohort study.

Methods: Cox proportional hazards method was used to calculate the risk of subsequent TB in individuals with CD. In a separate analysis, the risk of CD in individuals with prior TB was calculated using conditional logistic regression.

Results: CD was associated with an increased risk of subsequent TB (hazard ratio (HR) 3.74, 95% CI 2.14 to 6.53; p<0.001). Similar risk estimates were seen when the population was stratified for sex and age at CD diagnosis. Individuals with CD were also at increased risk of TB diagnosed in departments of pulmonary medicine, infectious diseases, paediatrics, or thoracic medicine (HR 4.76, 95% CI 2.23 to 10.16; p<0.001). The odds ratio for CD in individuals with prior TB was 2.50 (95% CI 1.75 to 3.55; p<0.001).

Conclusions: CD is associated with TB. This may be due to malabsorption and lack of vitamin D in persons with CD. Individuals with TB and gastrointestinal symptoms should be investigated for CD.
2003 through the Swedish National Inpatient Register (IPR). The IPR contains individual based data from hospital based discharge diagnoses in selected parts of Sweden since 1964 and has covered all of Sweden since 1987. Every record in the IPR can be identified through a personal identity number, which is a unique number assigned to more than 99.9% of all Swedish residents and immigrants.22

Individuals who received any of the following international classification of disease (ICD) codes between 1964 and 2003 were defined as having CD: ICD-7: 286.00; ICD-8: 269.00, 269.98; ICD-9: 579A; ICD-10: K09.0. For each individual with CD, Statistics Sweden used the national Total Population Register to select up to five reference individuals matched for age, sex, calendar year, and county. The Total Population Register includes information on area of residence, vital status, and dates of immigration or emigration.

The IPR was then used to identify individuals with TB among the study participants. TB was defined according to the following ICD codes: ICD-7: 001-019; ICD-8: 010-019; ICD-9: 010-018; ICD-10: A15-A19. The various codes represent different forms of M tuberculosis infection (for example, ICD-9: 010 is primary TB; 011 is lung TB; 012 is other forms of TB in the lungs including pleural TB; 013 is TB in the central nervous system, 014 is TB in the abdomen, 015 is TB in the skeleton or joints; 016 is urogenital TB; 017 is TB in other organs including the skin; and 018 is military TB). We specifically recorded individuals with gastrointestinal TB (any of the following ICD codes: ICD-7: 011; ICD-8: 014; ICD-9: 014; ICD-10: A18.3).

We also identified individuals with a hospital discharge diagnosis of sarcoidosis during the 3 year period after the first recorded diagnosis of TB. A diagnosis of sarcoidosis shortly after a TB diagnosis may signal misclassification and a false positive TB diagnosis. Sarcoidosis was defined according to the following ICD codes in the years 1964–2003: ICD-7: 138.00–138.10; ICD-8: 135; ICD-9: 135; ICD-10: D86.

Exclusion criteria
The Swedish National Board of Health identified 15 533 individuals with CD. Ninety four were excluded because of data irregularities such as death recorded before the first diagnosis of CD. One of the excluded individuals also had a diagnosis of TB. In order to minimise the risk of detection bias, 50 individuals were excluded because TB was diagnosed before study entry and a further 13 were excluded because TB was diagnosed within the first year after the initial CD diagnosis. A further 1040 individuals were excluded because they were followed for less than 1 year, and one individual was excluded because of probable misclassification (the hospital discharge diagnosis consisted partly of a TB ICD code and partly of another ICD code). Similar exclusion criteria were applied to reference individuals. The analyses of the current study were therefore based on 14 335 individuals with CD and 69 888 reference individuals who had never had CD. Both cohorts were free of TB at the start of the follow up period. Their characteristics are shown in Table 1.

One individual with TB had received a diagnosis of sarcoidosis during the 3 year period after the first recorded TB diagnosis. All three individuals with a diagnosis of TB where gastrointestinal TB had been recorded as the first recorded TB diagnosis were among those excluded for other reasons. The analyses therefore do not include any individual initially diagnosed with gastrointestinal TB.

A separate analysis of the risk of CD in individuals with prior TB consisted of 15 398 individuals with a later diagnosis of CD and 76 857 reference individuals without a later diagnosis of CD.

Socioeconomic index (SEI)
Socioeconomic data were available for a subset of 45 471 individuals (table 1) based on classification of occupation from 1968 into three categories where category I denotes a high SEI. About 6500 of these were children, and they were assigned a socioeconomic code on the basis of the occupation of their mother because divorce is not uncommon in Sweden and children then tend to live with their mothers.

Statistical methods and analyses
Cox regression was used to calculate hazard ratios (HRs) for subsequent TB in individuals with CD. The follow up time started 1 year after study entry (date of first recorded diagnosis of CD and corresponding date in reference individuals) and ended on the date of first discharge diagnosis of TB, date of emigration, death, or the end of the study period (31 December 2003), whichever came first.

The Cox model was internally stratified for age, sex, calendar year, and county, therefore resembling a conditional logistic regression as each individual with CD was only compared with his/her matched reference controls. This also means that a different baseline hazard function for the risk in the reference group is presumed for each stratum. This is necessary as the follow up begins at different ages/periods for the patients with CD and the matched comparison population, for whom the follow up begins at the same age period as the patient with CD in their risk set.

In a separate analysis we stratified for age at first recorded diagnosis of CD (<15 v ≥16 years) and sex. In order to increase the specificity of the diagnosis of TB, we specifically looked at individuals who had (1) received their first TB diagnosis in a department of pulmonary medicine, infectious diseases, paediatrics or thoracic medicine; (2) TB listed as the main/first diagnosis (rather than a secondary diagnosis); or (3) no diagnosis of sarcoidosis within 3 years of the first diagnosis.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Characteristics of participants in the main analysis of coeliac disease (CD) and subsequent tuberculosis (TB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristics</td>
<td>No CD (n=69888)</td>
</tr>
<tr>
<td>Age at first recorded diagnosis of CD (years)</td>
<td></td>
</tr>
<tr>
<td>0–15</td>
<td>–</td>
</tr>
<tr>
<td>16+</td>
<td>–</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>28724 (41.1)</td>
</tr>
<tr>
<td>Women</td>
<td>41164 (58.9)</td>
</tr>
<tr>
<td>Calendar period</td>
<td></td>
</tr>
<tr>
<td>1964–1973</td>
<td>2422 (3.5)</td>
</tr>
<tr>
<td>1974–1983</td>
<td>19174 (27.4)</td>
</tr>
<tr>
<td>1984–1993</td>
<td>30733 (44.0)</td>
</tr>
<tr>
<td>1994–2003</td>
<td>17559 (25.1)</td>
</tr>
<tr>
<td>Socioeconomic index*</td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>7278 (10.4)</td>
</tr>
<tr>
<td>II</td>
<td>9251 (13.2)</td>
</tr>
<tr>
<td>III</td>
<td>20134 (28.8)</td>
</tr>
<tr>
<td>Missing data</td>
<td>33225 (47.5)</td>
</tr>
</tbody>
</table>

Data presented as n (%).

Individuals with more than 1 year of follow up after diagnosis of CD or corresponding date in matched individuals (see text).

*Socioeconomic index: I is the highest category (see text). For reference individuals, the number of individuals who constituted the basis for the Cox regression is given. Data on socioeconomic index were available for another 6527 reference individuals but they were not part of the internally stratified calculations because of missing values on socioeconomic index in the matched individual with CD. When the 6527 reference individuals were added to those presented in the table, the proportion of missing values was similar among individuals with CD and those without CD.
of TB. Statistical significance was defined as 95% confidence intervals for HRs not including 1.00.

Conditional logistic regression estimated the risk of CD (the dependent variable) in individuals with prior TB. This analysis was performed in order to see whether or not the association between CD and TB was independent of sequence—that is, if there was only an association with TB after a diagnosis of CD and not before such diagnosis. The end of the follow up period was defined as the date of first diagnosis of CD and the same date for the matched controls without CD. Those with 1 year or less between the date of the first TB diagnosis and CD diagnosis and the corresponding date in controls were excluded. Risk estimates are given as odds ratios (OR). Statistics were calculated using SPSS 11.0 (Chicago, IL, USA).

Study power
The study had an 80% power to detect a 2.5-fold increased risk of TB in men and a 2.5-fold increased risk of TB in women. Overall, the study had an 80% power to detect a 2.0-fold risk increase for TB at a 5% significance level; this corresponds to 13 cases of subsequent TB among individuals with CD (some 14 000) compared with 35 cases among the referents (some 70 000).

Ethics
This study was approved by the Research Ethics Committee of the Karolinska Institutet. None of the participants was contacted. Patient information was anonymised prior to the analyses.

RESULTS
The median age at study entry was 3 years in both individuals with CD and those who had never had a diagnosis of CD (range 0–94 years). About 5000 individuals with CD had been diagnosed in adulthood. Most of the study participants were female (table 1). The median age at the first recorded diagnosis of TB was 66.5 years (range 1–87) in individuals with prior CD. The median time between the first CD diagnosis and the first recorded TB diagnosis was 6.5 years (range 1–26).

Among the 15 438 individuals with CD (all individuals with CD except those with data irregularities) there were 87 cases of TB before or after study entry (0.56%) compared with 35 cases among the referents (some 70 000)

CD and subsequent TB
TB was more common among patients with CD (HR 3.74, 95% CI 2.14 to 6.53; p<0.001, table 2). The increased risk was seen in both men and women (table 2). Five of 9368 children later developed TB. This corresponds to a threefold increase in the risk of TB in individuals who had received their first diagnosis of CD during childhood (HR 3.44, 95% CI 1.09 to 10.85; table 2). Individuals with CD were also at increased risk of TB diagnosed in departments of pulmonary medicine, infectious diseases, paediatrics, or thoracic medicine (HR 4.76, 95% CI 2.23 to 10.16; p<0.001; 31 positive events). Furthermore, TB listed as the main diagnosis was more common in individuals with CD than in reference individuals (HR 3.61, 95% CI 1.92 to 6.76; p<0.001; 47 positive events). Only one patient had received a diagnosis of sarcoidosis during the first 3 years after the first inpatient diagnosis of TB; exclusion of this patient from the analyses did not affect the risk estimates (HR 3.87, 95% CI 2.21 to 6.80; p<0.001). Data on SEI were available for a subset of individuals; adjustment for SEI did not affect the risk estimate for TB (crude HR 4.02, 95% CI 1.57 to 10.28; p = 0.004; adjusted HR 4.17, 95% CI 1.60 to 10.86; p = 0.003; 20 positive events).

Including the first year after study entry in the follow up, CD remained a risk factor for subsequent TB (HR 4.77, 95% CI 2.98 to 7.63; p<0.001; 81 positive events).

CD and prior history of TB
Conditional logistic regression showed an increased risk of CD in individuals with prior TB (OR 2.50, 95% CI 1.75 to 3.55; p<0.001).

DISCUSSION
The risk of TB in patients with CD was increased 3–4-fold and was also seen when stratified for age and sex. Our study confirms the results of the case-control study by Williams et al14 in 1988 and is consistent with the findings of a recent study by a US/Swedish research team.19 Unlike the latter study which showed a sixfold increase in death from TB among individuals with CD, the current study is not limited to deaths but is based on data from incident cases of TB. Williams et al found a history of TB in six of 76 individuals with CD.14 One of these patients had received a diagnosis of CD simultaneously with a diagnosis of TB, and in four patients the diagnosis of TB had preceded that of CD. In contrast, we chose to focus on the risk of future TB in a cohort of individuals with diagnosed CD, although we also looked at the risk of CD in individuals with prior TB.

We excluded the first year after study entry in order to minimise the risk of detection bias. In contrast to the study by Williams et al,14 we used a matched approach where each individual with CD was only compared with his/her reference individuals. This means that important potential confounding factors such as age, sex, calendar year, and geographical location within Sweden have little effect on our risk estimates. In our study there was a statistically significant increase in the

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Coeliac disease (CD) and risk of subsequent tuberculosis (TB) in subgroups of patients with CD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No of participants</td>
</tr>
<tr>
<td>No CD</td>
<td>69888</td>
</tr>
<tr>
<td>Any CD</td>
<td>14335</td>
</tr>
<tr>
<td>Age at first recorded diagnosis</td>
<td></td>
</tr>
<tr>
<td>&gt;0–15 years</td>
<td>9368</td>
</tr>
<tr>
<td>&gt;16 years</td>
<td>4967</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>5909</td>
</tr>
<tr>
<td>Female</td>
<td>8426</td>
</tr>
</tbody>
</table>

HR, hazard ratio. Estimates derived from Cox regression internally stratified for sex, age, year of study entry, and county. The HR for TB is increased for individuals with CD diagnosed during childhood although the number of events is low (five individuals with subsequent TB). This is so since each individual with a diagnosis of CD in childhood is compared only with his/her five matched reference individuals, also children at study entry.
risk of TB among both men and women, and in individuals diagnosed with CD in either childhood or adulthood.

In contrast to earlier research,16-18 25 our study also has the advantage of a very large sample of exposed individuals and some 60 individuals with a later diagnosis of TB. We used national register data of more than 14 000 individuals with CD and some 70 000 reference individuals. Our population based approach and the public healthcare system in Sweden minimise the risk of biased ascertainment due to socioeconomic inequalities. Although TB is strongly linked to socioeconomic factors,26 these factors do not appear to be of importance to a diagnosis of CD.27 Socioeconomic data were lacking for a number of individuals; however, such data were available for more than 8800 individuals with CD. Adjustment for SEI did not affect the risk estimate; this is important because SEI among Swedes is strongly linked to smoking habits.28

Smoking has repeatedly been shown to increase the risk of TB29 and also to influence the progress and care seeking behaviour at hospital.30 The IPR does not contain any smoking data so we could not adjust for smoking. However, smoking is unlikely to induce a positive association between CD and TB because it has, if any,31 a negative32 effect on CD but a positive effect on the risk of TB. This is not consistent with our findings of a positive association between CD and TB, but we cannot rule out the possibility that inclusion of smoking in our statistical models could have influenced the HRs. We also cannot exclude the possibility that our study suffers from ascertainment bias. Some individuals with gastrointestinal TB may have been screened for CD and only later received a diagnosis of TB. However, none of the individuals with TB in the prospective study (risk of TB in individuals with CD) had received a diagnosis indicating gastrointestinal TB.

One strength of the current register study is the high specificity of chronic diseases in the IPR, including CD33 and TB.34 The high specificity of CD partly stems from the well established practice in Sweden of performing a small bowel biopsy before confirming a diagnosis of CD. The Swedish IPR has previously also been used to identify cases of TB.35 In TB, diagnosis may involve a tuberculin skin test and can be confirmed by radiography and sputum examination. The IPR does not contain any data on the investigations carried out in conjunction with a certain diagnosis, nor does it contain any physical data such as body weight or body mass index. As individual identities are customarily removed from datasets delivered by the Swedish National Board of Health and Welfare before access by researchers, we were unable to verify the diagnoses of TB in this study against other sources of information. However, in a smaller Swedish study on tumour necrosis factor antagonists and risk of TB, all hospitalised TB cases (which occurred in the IPR) reflected true TB.36 We increased the specificity of our outcome measure by restricting TB cases to those diagnosed in departments that often handle cases of suspected TB. This yielded an HR of 4.76. Although there is a risk of misclassification of TB, this should not be a major concern in the current study. Only when misclassification is differential by nature will it affect risk ratios more than marginally. Differential misclassification of CD status is unlikely.

There is a risk that not all individuals with CD or TB are identified using a hospital based approach. Not all patients with TB are cared for as inpatients, and we do not claim to have complete ascertainment of cases in Sweden from 1964 to 2003. However, admission to hospital for TB may be more common in Sweden than, for example, in the US.37 In the study by Askling et al.,14 12 of 15 patients with TB identified from a source other than the IPR had at some stage been admitted to hospital because of their TB. Such hospital admissions may be for clinical investigation of symptoms and findings such as unexplained inflammation, fever and cough, but may also occur in conjunction with diagnostic procedures such as bronchoscopy. Comparative European data show that the incidence of TB in Sweden in 1998 was lower than in any of the other countries with which it was compared—for example, the incidence in Denmark and Finland was 10-12/100 000, in England and Wales 10-11/100 000, and in Germany 12-13/100 000.38 The lower incidence of TB and a relatively early decline in TB compared with other European countries may explain the small number of cases of TB in our study.

Since CD is not linked to the SEI, exposure to CD is unlikely to influence care seeking behaviour. However, since CD may differ between ethnic groups, we cannot exclude the possibility that ethnicity may have influenced our results. Although sensitivity may be an issue in this study, the numbers and the statistical power are large. It is important to note that the Swedish IPR also identifies individuals with CD when CD is listed as a secondary diagnosis and the main reason for hospitalisation is for a disease other than CD. Considering that the largest Swedish screening study for CD to date found a prevalence of diagnosed CD in adulthood slightly above 1/1000,39 and that the Swedish population is nine million people, we assume that a large proportion of individuals with CD have been identified in this study. It is common for individuals (especially children) to be admitted to hospital while first undergoing small bowel biopsy and then to be transferred to an outpatient setting. The frequent hospitalisation of children for small bowel biopsy partly explains the low median age at diagnosis of CD in the current study. The median age in our cohort is consistent with data from another large Swedish study of CD.40 False negative patients with CD are unlikely to influence the HRs in this study since they are outnumbered by healthy reference individuals.

There may be several explanations for the association between CD and TB. As a result of malabsorption of vitamin D and calcium in patients with CD,41 osteomalacia or osteopenia may occur.42 This mechanism may operate both before making a diagnosis of CD (when there is intense mucosal inflammation) and after such diagnosis. Individuals with a diagnosis of CD often have persistent low grade inflammation with affected vitamin status43 many years after the introduction of a gluten-free diet.44 In addition, the gluten-free diet and products are often low in vitamin D,45 increasing the risk of vitamin D deficiency in patients with CD. Because of its capacity to activate macrophages, vitamin D is important for generating an accurate immune response against TB. Vitamin D has been shown to induce the synthesis of nitric oxide in macrophages, thus acting to suppress the growth of M tuberculosis in these cells.46 Furthermore, vitamin D augments the effect of IFNγ in promoting the granulomatous process and promotes the differentiation of monocytes into epithelioid cells and multinucleated giant cells which form prominent parts of granulomas.47 Low levels of serum vitamin D have been found in patients with active TB.48 Furthermore, an association has been reported between genetic variants of the vitamin D receptor and TB infection.49 We therefore hypothesise that malnutrition in CD leads to malabsorption of a number of nutrients including vitamin D which increases the risk of TB infection. This underlines the importance of evaluating vitamin D status in patients with CD.

The strongest genetic links to CD are found within the MHC locus with the well established correlation to HLA-DQ2 (DQA1*05/DQB1*02).50 Associations between TB and various HLA alleles have been documented, although they are not as strong as in CD.51 To our knowledge, no other strong genetic associations common to CD and TB have been reported. From a
genetic standpoint, it is of interest that HLA-DQ in northern Europeans is often part of the so-called ancestral haplotype 8.1.7 This haplotype contains a number of genes including specific alleles of HLA class I and class II molecules, as well as genes for TNFα and complement factors C2 and C4. Since the C2 molecule, which is polymorphic, is important in the mycobacterial invasion of macrophages, it is possible that a particular C2 allele could promote TB infection in a subgroup of patients.

In some respects the mechanisms of cell mediated immune responses in CD and TB seem to be quite different. In CD the gluten-specific CD4+ T cells driving the pathology are of the Th1 type.8 In contrast, a robust Th1 response is needed for protection against M tuberculosis infection and rare genetic defects in the Th1 cytokine pathway (IL-12/IL-23 and IFNγ) increase the susceptibility to TB infection.9 Studies of cytokine levels in TB have given divergent results, but some studies have shown an increased Th2 profile.10 A polymorphism in the gene for TNFα has been found to have opposite associations with CD and TB, since the TNF-308A allele (promoting a high TNFα production) is a susceptibility factor for CD and autoimmune disease but a protective factor against TB.10

In conclusion, this study shows that CD is associated with an increased risk of TB. In countries where TB is uncommon, an increase in the relative risk of 3–4 translates into a rather small absolute increase. We hypothesise that the malnutrition associated with CD, particularly vitamin D deficiency, could play an important part in the increased risk of TB. As Peters et al.10 found a sixfold increase in the risk of death from TB in individuals with CD and we have reported a 3–4-fold increase in the risk of developing TB, we cannot exclude the possibility that CD may also be a complicating factor in existing cases of TB and may increase the severity of the disease.

Authors’ affiliations
J F Ludvigsson, Department of Paediatrics, Örebro University Hospital, Sweden
J F Ludvigsson, A Ekbom, S M Montgomery, Clinical Epidemiology Unit, Department of Medicine, Karolinska University Hospital/Institute, Sweden
J Wahlström, J Grunewold, Division of Respiratory Medicine, Department of Medicine, Karolinska University Hospital/Institute, Sweden
A Ekbom, Harvard Medical School, Boston, Massachusetts, USA
S M Montgomery, Clinical Research Centre, Örebro University Hospital, Sweden

This project was supported by a grant from the Swedish Society of Medicine, the Swedish Research Council, the Sven Jerring Foundation, the Örebro Society of Medicine, the Majblomman Foundation, and the Swedish Coeliac Society. JW and JG were sponsored by the Swedish Medical Research Council. JFL was sponsored by the Juhlin Foundation, the Cla Grösshoffs Foundation, and was supported by a grant from the Örebro University Hospital while writing this article.

Competing interests: none declared.

REFERENCES
β2 adrenoceptor polymorphisms (ADRB2 gene) are not strongly associated with asthma incidence or prevalence

The aim of this study was to analyse any potential correlation between β2 adrenoceptor (ADRB2) polymorphism and asthma risk and prognosis. The ADRB2 gene contains polymorphic variants with high minor allele frequencies in the white population. The three coding regions, Arg16Gly, Gln27Glu and Thr164Ile, are shown to be functionally relevant. Previous meta analyses looking into the correlation between the genetic polymorphism and asthma risk have produced conflicting results.

In this study people born in the UK during one week in March 1958 were analysed by means of interviews (inquiring after symptoms of asthma, wheezy bronchitis and wheezing) and lung function tests at various ages. Those who had a history of significant wheezy bronchitis during childhood were examined at age 34–35 years by trained nurses and spirometry was performed. People who were still in contact with the study at age 42–45 years had DNA extraction for genotyping, measurement of immunoglobulin E levels and spirometry. A separate group of 41 individuals with severe asthma was recruited for specific genotyping.

Half of the cohort of 8018 individuals had some history of wheezing by the age of 42. No significant association of any one single nucleotide polymorphism was seen with lifetime prevalence or age of asthma onset. However, there was a small correlation between persistence of asthmatic symptoms from childhood to middle age and Arg16Gly and Gln27Glu gene polymorphism. Childhood wheezers who were homozygotes for the Arg16-Gln27 haplotype had five or more wheezing episodes in the past year at the age 42. The authors also did a further meta analysis including the results of previous studies and this did not show a significant correlation between gene polymorphism and prevalence or severity of asthma.

The results of this genetic association study in a British population do not confirm a significant correlation between β2 adrenoceptor polymorphisms and asthma risk or prevalence. However, any potential correlation between such genetic findings and the prognosis of childhood wheezing needs to be determined with more specifically designed studies.

M Pagaria

Intermediate Trainee Intensive Care, Gloucestershire Royal Hospital, Gloucester, UK; dr_pagaria@yahoo.co.uk