
A.J. FOWELL1, P.W. THOMAS2, S.L. SURGENOR3 and J.A. SNOOK3

From the 1Division of Infection, Inflammation and Repair, University of Southampton, Southampton General Hospital, Southampton, and 2Dorset Research and Development Support Unit and 3Department of Gastroenterology, Poole Hospital, Poole, UK

Received 19 December 2005 and in revised form 24 March 2006

Summary

Background: Coeliac disease (CD) results from mucosal exposure to dietary gluten in genetically predisposed individuals, although other environmental factors may be involved. The seroprevalence of CD is approximately 1%, with a high ratio of undiagnosed to diagnosed cases, leading to the concept of a ‘coeliac iceberg’.

Aim: To provide contemporary estimates of the incidence of diagnosed CD and the size of the submerged ‘coeliac iceberg’, and to seek evidence of disease clustering.

Design: Prospective observational study in a defined local population.

Methods: Data were collected prospectively for all biopsy-proven cases diagnosed at Poole Hospital, 1993–2002. Age-specific incidence was calculated and point prevalence estimated for cases within the defined study zone. Evidence of disease clustering was sought using a space-time scan statistic based on a Poisson model.

Results: The overall incidence of CD was 8.7 cases/100 000/year (95%CI 7.4–10.1), with a median age at diagnosis of 53 years. Incidence increased progressively during the study period, and the estimated point prevalence of biopsy-proven CD rose from 0.18% to 0.4%. An area of significant space-time clustering was identified, with an incidence of 22.9 cases/100 000/year (95%CI 16.1–31.6), but there was no evidence of seasonality.

Discussion: The submerged component of the ‘coeliac iceberg’ may be diminishing due to increasing case ascertainment, with a projected ratio of undiagnosed to diagnosed cases as low as 1.5:1. Our identification of clustering must be interpreted with caution, but suggests that an additional environmental factor may influence the pathogenesis of CD.

Introduction

Coeliac disease (CD) is a chronic inflammatory disorder of the small bowel.1 A plethora of evidence implicates the immune system in the mediation of the inflammatory response in CD, including the strong link with HLA-DQ2 and HLA–DQ8 alleles,2 the association with other autoimmune disorders,3 and the frequent finding of anti-endomysial and anti-gliadin antibodies in patients with untreated CD.4–6

Population-based screening studies suggest an overall seroprevalence of CD in Western populations of about 1%.7 The clinical spectrum of CD...
is broad, with only a minority of patients presenting with classical features of weight loss and malabsorption. The recognition from screening studies that the majority of cases of CD are undiagnosed has led to the concept of a ‘coeliac iceberg’, with the submerged component representing the atypical and silent forms of the disease. The ratio of undiagnosed to diagnosed cases has been estimated to be as high as 6.5:1 in a UK population, making a case for unrestricted population screening.

There is overwhelming evidence that CD results from mucosal exposure to dietary gluten in genetically predisposed individuals. However, various observations (including the incomplete concordance for CD in monozygotic twins, and the relationship between cigarette smoking and the likelihood of developing CD) suggest that the aetiopathogenesis of CD cannot be fully accounted for on the basis of gluten exposure and genetic predisposition alone. The implication is that an additional environmental factor may be operative.

The demonstration of case-clustering of a disease supports the existence of an environmental trigger, since exposure (e.g. to an infective agent) may vary with time and location. Examination for clustering has previously been reported in other gastrointestinal disorders, but not in CD, although the possibility that a specific viral infection might predispose to the development of CD has been explored.

The aims of the current study were firstly to provide contemporary estimates of the incidence of diagnosed CD and the size of the submerged component of the ‘coeliac iceberg’, and secondly to seek evidence of disease clustering, in a geographically-defined population in East Dorset.

**Methods**

The study had the approval of the East Dorset LREC. Prospective data were collected for all biopsy-proven cases of CD diagnosed at Poole Hospital during the 10 years from 1993 to 2002. No formal population screening programme for CD was in place during this time, and all patients had been referred to the Gastroenterology department via standard primary and secondary care referral channels. Month of diagnosis, gender, age and postcode at diagnosis were recorded for each incident case.

To minimize the effects of hospital referral bias, a geographical study zone was defined, with margins perpendicular to lines joining Poole Hospital to the three nearest adjacent hospitals, the margins intersecting at points one third of the shortest distance between hospitals (Figure 1). Only cases residing within this zone at diagnosis were included in the study.

The UK Census 2001 was the source of detailed population data for census output areas (COAs) with a geographical centroid within the defined study zone. COAs are standardized geographical units of area generated by the Office of National Statistics (ONS) for the reporting of data from the 2001 Census. Individual COAs contained small numbers of cases of CD (maximum 3), with the majority containing none, and thus to highlight geographical variation, lower-layer super output areas (SOAs) were used. Lower-layer SOAs are a relatively recent geographical unit developed by the ONS. They comprise groups of COAs (typically five) with a minimum population of 1000 and a mean of about 1500, merged by a computer programme taking into account measures of population size, mutual proximity and social homogeneity.

Incidence data was plotted geographically using standard commercially available mapping software (Customised Mapping Limited, Yarnscombe, Devon, UK). The binomial test was used to assess whether the sex ratio amongst cases differed from 1:1. The incidence of coeliac disease was compared between age groups and over the course of the study using the \( \chi^2 \) test for association, and the \( \chi^2 \) goodness-of-fit statistic was used to compare the number of cases diagnosed per month with that expected from the length of the month. The analysis was conducted using SPSS for Windows v. 12, and a 5% significance level was used.

Statistical evidence of disease clustering was sought using spatial, temporal and space-time scan statistics based on a Poisson model, using SatScan software. From incident data for each of 628 COAs in each of the 10 years of the study, the scan statistic was used to identify groups of adjacent COAs and/or years in which the incidence of disease was higher than expected, given their population density and age distribution. Age (coded as <25 years, 25–44 years, 45–64 years, and ≥65 years) was included as a covariate in the statistical model. Observed and expected numbers of cases were incorporated into a likelihood function for each ‘scanning window’ of adjacent COAs and years, and this was then maximized to identify the most likely space-time cluster. The likelihood ratio for this window is the maximum likelihood test statistic, and a \( p \) value was calculated using Monte Carlo hypothesis testing. This method helps to ensure that spurious clusters are not identified as a result of multiple significance testing.
Results

During the study period, 159 of 220 incident cases of CD diagnosed at Poole Hospital resided within the study zone. There were more females than males (F:M ratio 2.1:1, \(p<0.001\)), and age at diagnosis was a median 53 years (IQR 26) and a mean 51.7 years (SD 19.0). The 628 COAs within the study zone provided a total population of 183,699, giving an overall CD incidence of 8.7 (95%CI 7.4–10.1) cases per 100,000 per annum. As shown in Figure 2, the age-specific incidence ranged from 2.5 per 100,000 per annum for the 0–14-year age group to 16.8 per 100,000 per annum for the 60–74 year age-group (\(\chi^2 = 49.5, \text{df} = 5, p<0.001\)).

There was a significant increase in the incidence of diagnosed CD during the study period (Figure 3), from 6.0 cases per 100,000 per annum in 1993–4 to 13.3 cases per 100,000 per annum in 2001–2 (\(\chi^2 = 18.0, \text{df}=4, p=0.001\)). Figure 4 shows the distribution of cases by month of diagnosis: there was no significant variation (\(\chi^2 = 12.9, \text{df}=11, p=0.30\)).

Space-time variation in the incidence of CD was observed within the study zone. SatScan analysis revealed an area of space-time clustering consisting of 110 COAs in the Wimborne Minster area (population size 32,308), located in the northeastern sector of the study zone (Figures 1 and 5).
During 1998–2002, this area yielded 37 cases of coeliac disease compared to an expected figure of 14.1, taking into account population density and age distribution ($p=0.009$). The incidence of CD in the ‘Wimborne cluster’ was 22.9 cases per 100,000 pa, and this cluster of 23 extra cases accounted for about half of the increased incidence of CD in the whole study zone, which rose from 56 cases in 1993–1997 to 103 cases in 1998–2003. No other statistically significant clusters were identified.

**Discussion**

We have calculated the age-specific incidence of diagnosed CD in a strictly-defined local area based on population data from the UK census. The results confirm that CD presents more frequently in adulthood than in childhood, and is more common in females.\(^{21-23}\) The overall incidence of 8.7 cases per 100,000 per annum in the Poole area is some five-fold greater than in studies reported prior to 1995.\(^{21-23}\) Publications from the UK and elsewhere confirm an increasing incidence of diagnosed CD over the last decade, with contemporary figures comparable with our findings.\(^{24-27}\) The results of studies of the incidence of clinically diagnosed CD are summarized in Table 1.

Enhanced case ascertainment is probably a major reason for the rising incidence of clinically diagnosed CD over recent years. Contributory factors are likely to have been increasing awareness of the myriad of presentations of CD, and the targeted screening of at-risk subjects by means of serology or duodenal biopsy. Examples of at-risk subjects include close relatives of index cases, and those with unexplained haematocrit deficiency or non-specific abdominal symptoms. The anti-endomysial antibody assay was available in Poole from 1996 (prior to which antigliadin antibody was the preferred serological test). It is likely that increased awareness and use of this more sensitive and specific assay contributed to the increased incidence of CD observed in the second half of the study period.

From the calculated incidence, an assumption that this incidence is at steady state, and a figure for mean life expectancy from diagnosis, one can estimate point prevalence. Given the mean age at diagnosis in this study of 51.7 years, and the evidence that the diagnosis of CD has no more than a marginal effect upon mortality,\(^{28}\) we have assumed a mean life expectancy from diagnosis of 30 years.\(^{29}\) Derived estimates of projected point prevalence of diagnosed CD for various incidence rates pertinent to this study are shown in Table 2. These estimates assume that the given incidence rate is sustained.

The overall population prevalence of CD reflects the arithmetic sum of diagnosed and undiagnosed cases. Serological studies of the overall prevalence of CD show some variation, but most indicate a seroprevalence of around 1% in Western populations in general\(^7\) and the UK in particular,\(^{30,31}\) and there is no convincing evidence that this has changed appreciably over the last decade. If the overall prevalence is constant and the prevalence of diagnosed CD rises, one might anticipate that the ratio of undiagnosed to diagnosed cases—and thus the size of the submerged component of the ‘coeliac iceberg’—would fall.

Previous reports have suggested a ratio of undiagnosed to diagnosed CD in the UK of 6.5:1.
Figure 5. Incidence of CD by lower-layer SOA for the full 10 years of the study (urban areas shaded).
Assuming an overall prevalence of 1% for CD in the study zone, our estimates suggest that this ratio in East Dorset might have fallen from 4.6:1 to 1.5:1 as the result of the rising incidence of diagnosed CD during the course of our observation period. Clearly these are projected figures, and only apply once a steady-state situation is reached. Nevertheless, a reduction in the size of the submerged component of the 'coeliac iceberg' weakens the argument for unrestricted screening of general populations for CD. Furthermore, the ratio is likely to be especially low in the older age groups, since the cumulative prevalence of diagnosed CD (unlike the seroprevalence) rises progressively with age.

Geographical variation of incidence might be expected with any number of diseases, as a result of geographic variations in population demographics and socio-economic factors. Space-time aggregation however, in which the disease incidence varies with both place and time, may be a more intriguing epidemiological finding, because explanations for clustering are limited to a restricted set of factors that vary in the same specific pattern as the disease. Clustering of disease in space and time is, in theory at least, likely to relate to environmental factors that move from place to place or suddenly appear in specific locations.

There are several reasons for supposing that environmental triggers other than gluten may contribute to the aetiopathogenesis of CD. Firstly, if genetic factors and gluten exposure were the sole determinants, all predisposed individuals might be expected to develop CD after first exposure in early childhood; yet the median age at diagnosis in this study was 53 years, and the majority of new cases of CD present in adulthood. Secondly, there are well-documented reports of subjects with histologically normal small-bowel mucosa, who have gone on to develop CD later in life. Thirdly, only a minority of genetically predisposed individuals who are exposed to gluten progress to CD, with incomplete concordance for CD in monozygotic twins. Finally, a series of recent reports have demonstrated that cigarette smoking in adults may, for unclear reasons, substantially reduce the likelihood of being diagnosed with CD. The implication of this collection of observations is that an additional environmental factor, as yet undefined, may be operative.

We have identified a statistically significant space-time cluster of CD in the Wimborne Minster area of Dorset that is not accounted for by the age distribution of the local population. Wimborne is a small market town in a rural area of East Dorset, with some supporting light industry. In keeping with the epidemiology of CD, it is of interest to note that the Wimborne area has a number of light industry establishments that manufacture gluten-free products.
with much of Dorset, the area is moderately affluent, with (compared to national figures) a relatively low population density and proportion of younger people. There was little change in the size of the local population between the 1991 and 2001 censuses.

Before reading too much into the finding of a cluster, a number of important potential confounders must be considered. Firstly, CD is commoner in females, and has a tendency to aggregate in families. Data on gender is available from the UK census, although these data are not broken down by age, and therefore could not be introduced into the SatScan analysis as independent co-variates. However, the sex ratio within the cluster area was no different to that in the study area as a whole (data not shown). Only 12% of our subjects had a family history, and this is unlikely therefore to be of relevance to the findings. Neither of these potential confounders would be expected to cause clustering in time.

Secondly, cigarette smoking is strongly associated with the likelihood of being diagnosed with CD.\(^{11–13}\) Data on smoking are not available from the UK census or other sources at sufficient geographical resolution for the purposes of this study. However, unpublished data from the largest General Practice in Wimborne indicates that the prevalence of cigarette smoking is little different from that of the local population.\(^36\)

Thirdly, ascertainment bias at the primary-care level might influence the geographical distribution of incident cases. We think that this is unlikely because: (i) the cluster is spread across a number of different individual primary care physicians and primary-care practice areas; (ii) unpublished studies on the incidence of other gastrointestinal disorders (ulcerative colitis and Crohn’s disease) have shown no excess in the Wimborne area; and (iii) there is no apparent excess of laboratory EMA requests from practices within the cluster area (Wimborne GPs care for roughly 17% of the population, and account for approximately 14% of requests for EMA from practices within the study zone).

Finally, we have attempted to minimize referral bias by defining a strict study zone around the base hospital. In semi-rural areas, geography (in particular, distance to the nearest hospital) is often a major determinant of referral practice to secondary care. Therefore the geographical location of the potential cluster on the edge of this area makes referral bias an unlikely explanation, since one might expect a lower number of cases preferentially referred to Poole from areas further away from the hospital.

If clustering is not easily explained by bias or confounding, we need to consider whether it might be due to environmental factors. Clearly there are a wide range of possible explanations, ranging from toxins to infective agents, and few specific clues. It is tempting to speculate that the clustering observed might reflect a viral trigger such as adenovirus 12,\(^17\) although specific supporting evidence for this particular hypothesis is lacking. It remains to be established whether the putative environmental factor is acting to increase the true incidence of CD or in some unspecified way to enhance case ascertainment. Further studies are planned to explore these possibilities.

Our data provide a ‘snapshot’ of the current epidemiology of CD in the UK, and argue against the need for generalized population screening in the light of improving case ascertainment in clinical practice. Our preliminary evidence of space-time clustering of CD requires both confirmation and further investigation.

### Acknowledgements

We are indebted to David Yelland of Customised Mapping Ltd, Yarnscombe, Devon, UK for the development of the incidence maps. We are grateful to Dr Nicholas Sharer for his support and constructive criticism, and to Dr Paul Massey of the Department of Immunology, Royal Bournemouth Hospital, for data regarding EMA requests.

### References


20. Kulldorff M. Information Management Services, Inc. SatScan v.3.0.2. Software for the spatial and space-time scan statistics. Bethesda MD, National Cancer Institute, 2002.


