Infections and Risk of Celiac Disease in Childhood: A Prospective Nationwide Cohort Study

Karl Mårild, MD, PhD[1,2], Christian R. Kahrs, MD[1,3], German Tapia, PhD[1], Lars C. Stene, PhD[1] and Ketil Størdal, MD, PhD[1,3]

OBJECTIVES: **Studies on early life infections and risk of later celiac disease (CD) are inconsistent but have mostly been limited to retrospective designs, inpatient data, or insufficient statistical power. We aimed to test whether early life infections are associated with increased risk of later CD using prospective population-based data.**

METHODS: **This study, based on the Norwegian Mother and Child Cohort Study, includes prospective, repeated assessments of parent-reported infectious disease data up to 18 months of age for 72,921 children born between 2000 and 2009. CD was identified through parental questionnaires and the Norwegian Patient Registry. Logistic regression was used to estimate odds ratios adjusted for child's age and sex (aOR).**

RESULTS: **During a median follow-up period of 8.5 years (range, 4.5–14.5), 581 children (0.8%) were diagnosed with CD. Children with ≥10 infections (≥fourth quartile) up to age 18 months had a significantly higher risk of later CD, as compared with children with ≤4 infections (≤first quartile; aOR=1.32; 95% confidence interval (CI)=1.06–1.65; per increase in infectious episodes, aOR=1.03; 95% CI=1.02–1.05). The aORs per increase in specific types of infections were as follows: upper respiratory tract infections: 1.03 (95% CI=1.02–1.05); lower respiratory tract infections: 1.12 (95% CI=1.01–1.23); and gastroenteritis: 1.05 (95% CI=0.99–1.11). Additional adjustments for maternal CD, education level, smoking, birth weight, prematurity, infant feeding practices, birth season, and antibiotic treatment yielded largely unchanged results.**

CONCLUSIONS: **This is the first large-scale population-based cohort study of this association. Our results are in line with immunological data suggesting that early life infections may have a role in CD development. However, non-causal explanations for this association due to surveillance bias and reverse causation cannot be excluded.**

SUPPLEMENTARY MATERIAL is linked to the online version of the paper at http://www.nature.com/ajg

Am J Gastroenterol advance online publication, 8 September 2015; doi:10.1038/ajg.2015.287

INTRODUCTION

Celiac disease (CD) is a life-long immune-mediated disease prevalent in 1–2% of Caucasian populations (1). Patients with CD, including children, suffer a modestly increased mortality risk (2). In CD, dietary gluten causes villous atrophy and inflammation of the small intestine. The etiology of CD is considered to be multifactorial, where genetic and environmental factors interplay in triggering the disease (3). However, little is known about the nature and timing of such environmental factors.

Only one earlier general population-based cohort study has addressed the risk of CD following early life infections. That study, including 44 children with CD, did not find a significant association between parent-reported infectious disease in the child's first year of life and later CD (4). Findings from other studies are inconsistent (5–10), but have mostly been limited by retrospective designs (7) (liable to recall bias) and restriction to inpatient infectious disease data (5,8–10) and are therefore more susceptible to selection bias, or lacking data on potential confounders (e.g., breastfeeding and birth season).

[1]Division of Epidemiology, Norwegian Institute of Public Health, Oslo, Norway; [2]Department of Medical Epidemiology & Biostatistics, Karolinska Institutet, Stockholm, Sweden; [3]Department of Pediatrics, Østfold Hospital Trust, Fredrikstad, Norway. **Correspondence:** Karl Mårild, MD, PhD, Division of Epidemiology, Norwegian Institute of Public Health, PO Box 4404 Nydalen, Oslo 0403, Norway. E-mail: karlmarild@gmail.com
Received 5 May 2015; accepted 20 July 2015

There are many potential mechanisms in which infections may have a role in CD development. Viruses may induce pro-inflammatory interferons and cause upregulation and release of tissue transglutaminase, an enzyme important in enhancing the immunogenicity of gluten (11,12). Infections may also increase the intestinal epithelial permeability enabling epithelial translocation of gluten, a key process early in CD pathogenesis (13).

In this study, we used prospective cohort data to examine the association between early life infections and later CD. Given the ubiquitous nature of infections, we hypothesized that a high infection frequency in early life would be associated with later CD.

METHODS

This study is part of The Norwegian Mother and Child Cohort Study (MoBa), a prospective population-based pregnancy cohort study conducted by the Norwegian Institute of Public Health (14). Pregnant women were recruited from all over Norway in the years 1999–2008, and 40.6% of eligible women participated. The cohort includes 106,890 live-born children and 95,200 mothers. In MoBa, questionnaires were administered from early pregnancy and at regular intervals during childhood. For the purpose of this study, we used questionnaire data filled out around weeks 18 and 30 of pregnancy and at child's age of 6 months, 18 months, and 7–8 years. Questionnaires are available at www.fhi.no/moba. The current study used version VIII of the quality-assured data files released for research in February 2014.

Study population

Out of the 106,890 live-born children participating in MoBa, the parents of 87,973 children provided infectious disease data up to child's age of 6 months. In 72,921 children infectious disease data were also assessed in the 18-month questionnaire (see flow chart of study participants, **Figure 1**).

Outcome: CD diagnosis

Children diagnosed with CD by the end of 2013 were identified through combined data from the Norwegian Patient Registry (NPR) and parental questionnaires. The NPR contains prospectively recorded inpatient and hospital-based outpatient diagnostic data since 1 January 2008. Reporting to the NPR is mandatory and linked to the governmental reimbursement system for funding of health services, enabling a close to complete coverage of hospital-based pediatric health care, which is responsible for the diagnostic work-up for CD in Norway. Although diagnostic data from pediatricians with private practices were not included in the NPR data file used for this study, previous data suggest that in Norway pediatric CD is almost exclusively diagnosed by publicly funded pediatric caregivers (15). Linkage of MoBa-data to the NPR was possible using the unique personal identification number assigned to all Norwegian citizens. To reduce the possibility of including false-positive cases, we defined CD as at least two records of the International Classification of Diseases-10 code K90.0. This is because children undergoing CD investigation may sometimes receive a preliminary working diagnosis of CD while waiting for histological confirmation of the disease.

Children diagnosed prior to the launch of NPR (January 2008), and not followed as recommended in the outpatient clinic, may not be identified in the NPR. We therefore collected data on CD from parental questionnaires administered at child's age of 7–8 years. The MoBa-children were not screened for CD, and hence in this study CD represents clinically diagnosed disease.

Among the 87,973 children with infectious disease data up to the age of 6 months, 674 were diagnosed with CD by the end of follow-up (31 December 2013); of the 72,921 children with available infectious disease data up to the age of 18 months, 581 had CD. Children with unconfirmed CD (i.e., a single entry in the NPR without questionnaire confirmation; n=139) were excluded from the study (**Figure 1**). In a recent validation of CD in the MoBa-cohort, >92% of the CD diagnoses were confirmed by the parents and >99% of the diagnoses had been based on biopsy and/or a positive CD serology (a detailed account of the validation has been described elsewhere (16)).

Main exposure: parent-reported infection frequency in early life

Parent-reported infectious disease data were collected from questionnaires at age 6 and 18 months, including data on the type and number of infectious episodes the child had experienced up to the corresponding age. Type of infection was specified as common cold, throat infection, ear infection, croup, lower respiratory tract infection (stated as pneumonia or bronchitis), urinary tract infection, and gastroenteritis. We examined the total number of infections in the child, representing all parent-reported infectious episodes stated above, as well as the specific number of upper respiratory tract infections (i.e., common cold, throat infection, ear infection and croup), lower respiratory tract infections, and gastroenteritis. To reduce the impact of erroneously registered data, children reported to have ≥10 episodes of a specific infectious type over a 6-month period (e.g., ≥10 ear infections at age 12–18 months) were classified as missing in relevant analyses.

To assess infection severity, the parents reported infections requiring medical care (stated as an infection requiring a doctor's visit and/or hospital admission in the first 6 months or hospital admission for infection at an age of 6–18 months). Infections requiring medical care were categorized into any infection, upper or lower respiratory tract infection and gastroenteritis (see description above).

Our main analysis addressed the risk of later CD according to the overall infection frequency in the first 18 months of life, whereas secondary analyses considered specific type of infections: upper respiratory tract infections, lower respiratory tract infections, and gastroenteritis.

Potential confounders

Using data from pregnancy questionnaires we adjusted for maternal smoking in pregnancy (non-smoker, occasional, or daily smoker) and maternal education level (≤12, 13–15 or ≥16 years of education), as these factors have in previous studies been independently associated with early life infections (17) and, although

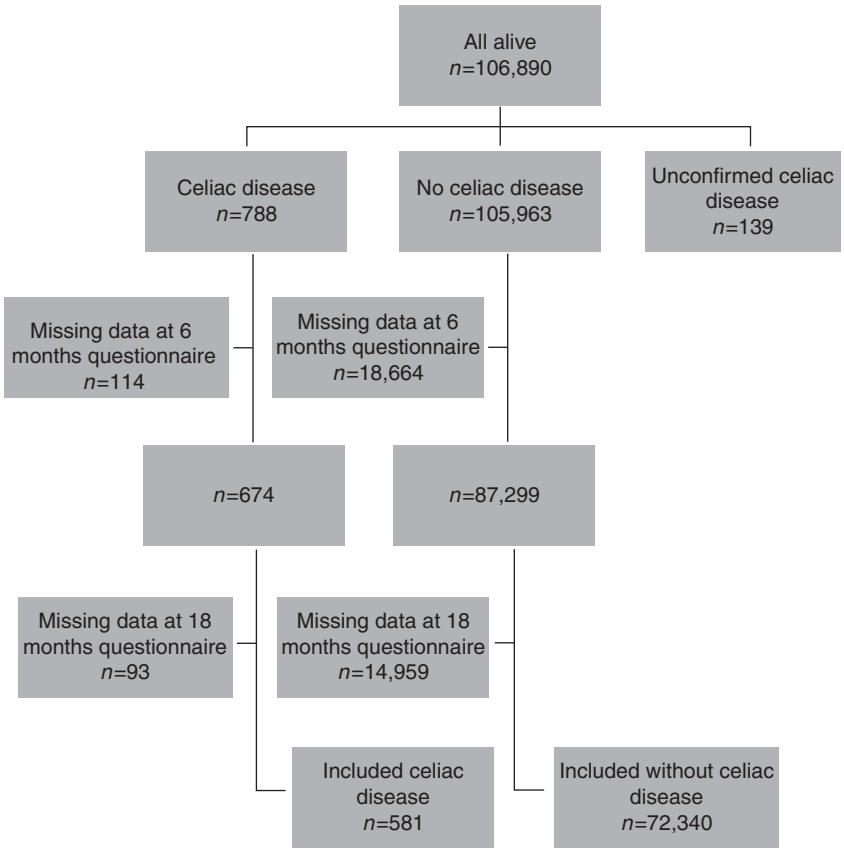

Figure 1. Flow chart of study participants. Children with a single record of celiac disease in the Norwegian Patient Registry (without questionnaire confirmation) were regarded to have unconfirmed celiac disease and excluded from the study.

much less clear, with offspring CD (8,16,18). Pregnancy questionnaires provided information on maternal CD, whereas data on age at the time of gluten introduction (<6 or ≥6 months) were collected from the 6-month questionnaire.

Breastfeeding has been associated with a lower risk of early life infections (19) and may, despite contradictory results (20,21), also be associated with pediatric CD. Using questionnaire data collected at age 6 and 18 months, we adjusted for duration of breastfeeding categorized into the following: no breastfeeding; <4 months; 4–5.9 months; 6–8.9 months; 9–11.9 months, or ≥12 months. We also defined antibiotic treatment as a potential confounder (22). Data on the use of systemic antibiotics (yes/no) up to 6 months or at an age of 6–18 months were collected from the 6- and 18-month questionnaire, respectively.

From the Medical Birth Register of Norway (23) we retrieved data on birth weight (<2,500 g; 2,500–3,499 g; 3,500–4,499 g and ≥4,500 g), parity ('0' (first child); 1; 2 or ≥3), and preterm birth (<37 gestational weeks), as these variables may be associated with both an increased risk of early life infections (17) and with offspring CD (5,8).

Sensitivity analyses

Age at CD diagnosis. To avoid potential recall bias or reverse causality, we performed a sensitivity analysis excluding children diagnosed with CD ≤18 months of age or those where age at the time of diagnosis could not be confirmed. Data on age at diagnosis were retrieved from a CD validation questionnaire collected in year 2014 and from the NPR, in which first recorded day ward admission for CD was regarded as the time of diagnosis. Procedure codes are not listed in the NPR. Out of the 581 children with CD and infectious disease data up to an age of 18 months, 128 were excluded because of missing data on age at the time of diagnosis (biopsied before the start of NPR or lacking data from validation questionnaire), and 21 children were excluded because of diagnosis ≤18 months of age (range, 10–18 months). Hence, 432 children with CD diagnosed >18 months of age were included in these sensitivity analyses.

Season of birth. Most (24–26), but not all (6,9), studies have found a positive association between CD and summer birth, as compared with winter birth. This association has been hypothesized to be due to gluten introduction in the winter when viral infections are more frequent. In a first test of this hypothesis, we examined the association between birth season and CD, adjusting for infection frequency and age at gluten introduction. In a separate analysis, we introduced birth season as a covariate in the analyses on infection frequency and later CD. Birth season was categorized accordingly: winter (December–February; reference

category), spring (March–May), summer (June–August), and autumn (September–November).

Statistical analyses

We used logistic regression to estimate odds ratios (ORs) as measures of the relative risk of CD. Robust variance estimator(27) was used in the main analysis given potentially cluster-correlated data among siblings. Infection frequency per type of infection was, given its distribution in the cohort, categorized into quartiles, tertiles, or dichotomous variables (with lowest category as reference). Because of the discrete nature of infection frequency (i.e., can assume only a countable number of values), some variables have a somewhat uneven distribution of individuals between the given quartiles/tertiles. We estimated ORs per increase in the number of infections as a test for trend. Potential deviation from linearity was examined by testing models with squared terms of the independent variable.

Infections after 6 months of age (a time when gluten often is introduced) may have a different impact on the risk of later CD, as compared with infections before 6 months of age. In pre-planned subanalyses, we therefore examined the risk of CD according to both infection frequency and infections requiring medical care in the first 6 months or at age 6–18 months.

All analyses were adjusted for child's attained age at the end of follow-up and sex (model I). In model II, we also adjusted for maternal CD, maternal smoking in pregnancy, maternal education level, parity, birth weight, prematurity, age at the time of gluten introduction, and duration of breastfeeding. In a separate adjustment model, we also chose, in addition to age and sex, to adjust for parent-reported use of systemic antibiotics up to age of 18 months.

Statistical significance was defined as 95% confidence intervals (CIs) for ORs not including 1.00. We used SPSS version 22.0 (IBM SPSS, Chicago, IL) and Stata version 13.0 (Stata, College Station, TX) for the statistical analyses.

Post hoc analysis

In a *post hoc* analysis, we examined whether the child's adherence to the Norwegian immunization schedule (28) up to age 18 months influenced the association between early life infection frequency and later CD. Parent-reported data on the child's adherence to the Norwegian immunization scheduled (complete vs. incomplete adherence) were collected from questionnaires at an age of 6 and 18 months and introduced as a covariate in the analysis on infection frequency and later CD.

Ethics

This study was approved by The Regional Committee for Medical Research Ethics in South–Eastern Norway. Written informed consent was obtained from all study participants. The MoBa study has obtained a license from the Norwegian Data Inspectorate.

RESULTS

Most children with CD were girls and the median age at the end of follow-up was 8.5 years (**Table 1**). Approximately 3% of the children with CD had a history of maternal CD compared with 0.4% among children without CD. In total, we had data on 564,516 infectious episodes with a median number of seven infectious episodes per child up to the age of 18 months (**Table 1**, **Supplementary Table S1** online).

Infection frequency

Children with ≥10 infections (corresponding to ≥fourth quartile) up to age 18 months had a significantly higher risk of later CD, as compared with children with ≤4 infections (≤first quartile; adjusted OR; aOR=1.32; 95% CI=1.06–1.65; per increase in infectious episodes, aOR=1.03; 95% CI=1.02–1.05; **Figure 2**). The aORs per increase in specific infectious types in the first 18 months were as follows: upper respiratory tract infections: 1.03 (95% CI=1.02–1.05); lower respiratory tract infections: 1.12 (95% CI=1.01–1.23); and gastroenteritis: 1.05 (95% CI=0.99–1.11; **Table 2**). Additional adjustments for maternal CD, maternal smoking in pregnancy, education level, parity, birth weight, preterm birth and infant feeding practices (adjustment model II; **Table 2**), or antibiotic treatment in the first 18 months of life (**Supplementary Table S2**) yielded largely unchanged aORs. Repeating the analyses using the robust variance estimator gave similar 95% CIs. In addition, restricting our data to children aged at least 6 years at end of study (CD, *n*=572; without CD, *n*=70,221) gave largely unchanged aORs as compared with the analyses on the whole cohort (data not shown).

Similar results were seen in subanalyses on high infection frequency (≥fourth quartile vs. ≤first quartile) before 6 months of age (aOR=1.28; 95% CI=1.02–1.59) and at age 6–18 months (aOR=1.37; 95% CI=1.08–1.74). There were nonsignificant increased aORs for CD according to the exposure to gastroenteritis and upper and lower respiratory tract infections in the first 6 months of life (**Supplementary Table S3**). Children experiencing ≥7 upper respiratory tract infections (≥fourth quartile) at age 6–18 months had an aOR of 1.42 (95% CI=1.13–1.80) for later CD, as compared with children with ≤2 upper respiratory tract infections (≤first quartile). A similar significantly increased aOR for CD was found after exposure to lower respiratory tract infection at age 6–18 months (**Supplementary Table S4**).

Severity of infection

Table 3 presents the aORs for later CD according to infections requiring medical care in the first 18 months of life. Overall, any infection requiring medical care (aOR=1.25; 95% CI=1.06–1.48), and specifically for upper or lower respiratory tract infection in the first 18 months of life, was significantly associated with later CD (**Table 3**). Seeking medical care for gastroenteritis up to age 18 months was not significantly associated with later CD. Looking specifically at the first 6 months of life only medical care for lower respiratory tract infection was significantly associated with later CD (aOR=1.48; 95% CI=1.09–2.02) (**Supplementary Table S5**). The aORs for CD according to infections requiring medical care between 6 and 18 months of age were broadly consistent with the results above (**Supplementary Table S6**).

Table 1. Characteristics of children with and without celiac disease with available infectious disease data up to 6 months and 18 months of age, respectively

	Data up to 6 months of age		Data up to 18 months of age	
	Celiac disease	No celiac disease	Celiac disease	No celiac disease
Total	674	87,299	581	72,340
Age at end of 2013, years (median, range)	8.5 (4.5–13.5)	8.5 (4.5–14.5)	8.5 (4.5–13.5)	8.5 (4.5–14.5)
Girls, *n* (%)	413 (61.3)	42,528 (48.7)	352 (60.6)	35,283 (48.8)
Parity, n *(%)*				
0[a]	303 (45.0)	39,801 (45.6)	269 (46.3)	33,758 (46.7)
1	255 (37.8)	30,779 (35.3)	214 (36.8)	25,006 (34.6)
2	90 (13.4)	13,139 (15.1)	75 (12.9)	10,664 (14.7)
≥3	26 (3.9)	3580 (4.1)	23 (4.0)	2912 (4.0)
Maternal celiac disease[b], *n* (%)	23 (3.4)	328 (0.4)	21 (3.6)	259 (0.4)
Maternal education level[c], n *(%)*				
≤12 years	207 (30.9)	27,612 (32.0)	162 (28.1)	21,974 (30.6)
13–15 years	275 (41.1)	36,487 (42.3)	243 (42.1)	30,852 (43.0)
≥16 years	187 (28.0)	22,130 (25.7)	172 (29.8)	18,942 (26.4)
Total number of infectious episodes[d]				
Median (range)	1 (0–10)[e]	1 (0–22)[e,f]	7 (0–31)[g]	7 (0–58)[g,h]
Mean±s.d.	1.66±1.57	1.59±1.58	8.4±5.12	7.7±4.64

[a]'0' indicates no previous children before the birth of the child included in this study.

[b]Data from pregnancy questionnaires. Missing data in mothers of 1083 children and 582 children with infectious disease data up to 6 months and up to 18 months of age, respectively.

[c]Maternal education level at the time of pregnancy. Missing data in mothers of 1,075 children and 576 children with infectious disease data up to 6 months and up to 18 months of age, respectively.

[d]Includes parent-reported gastroenteritis, upper or lower respiratory tract infection, and urinary tract infection.

[e]*P*-value=0.11 for test of difference in distribution of the number of infections (Mann–Whitney *U* test).

[f]Twelve children were reported to have a total number of ≥14 infections before age 6 months, which by large constituted of frequent upper respiratory tract infections. Children reported to have ≥10 episodes of a specific infectious type in the first 6 months of life were classified as missing.

[g]*P*-value=0.01 for test of difference in distribution of the number of infections (Mann–Whitney *U* test).

[h]Nine children were reported to have a total number of >40 infections up to age 18 months, which by large were comprised of frequent upper respiratory tract infections. To reduce the impact of possibly erroneously registered data, children reported to have ≥10 episodes of a specific infectious type over a 6-month period were classified as missing (see Methods section).

Sensitivity analyses

Age at CD diagnosis. Restricting our analyses on infection frequency to the 432 children diagnosed with CD after age 18 months (excluding children with missing data (*n*=128) or diagnosed ≤18 months of age (*n*=21)), we found largely unchanged aORs as compared with the analyses on the whole cohort (**Supplementary Table S7**).

Season of birth. There was no significant association between birth season and CD (Likelihood ratio test, *P*=0.75) and aORs did not change substantially after adjustments for infection frequency and age at gluten introduction (**Supplementary Table S8**). Birth season had only a minor influence on infection frequency up to age 18 months (**Supplementary Table S9**). Introducing birth season as a covariate to the analysis on infection frequency and risk of later CD yielded only marginally changed aORs (data not shown).

Post hoc analysis

Out of the 71,061 children with data on vaccination coverage up to age 18 months, the parents of 67,942 children (95.6%) reported complete adherence to the recommended Norwegian immunization schedule (28). Including immunization status (complete vs. incomplete adherence) as a covariate, in addition to age and sex, yielded largely unchanged aOR for CD (e.g., per increase in infectious episodes up to age 18 months, aOR=1.03; 95% CI=1.02–1.05).

DISCUSSION

In this large-scale prospective cohort study, we found a positive association between parent-reported infections the first 18 months of life and later CD. These associations remained significantly increased after adjustment for potential confounders, e.g., antibiotic treatment. Although repeated respiratory tract

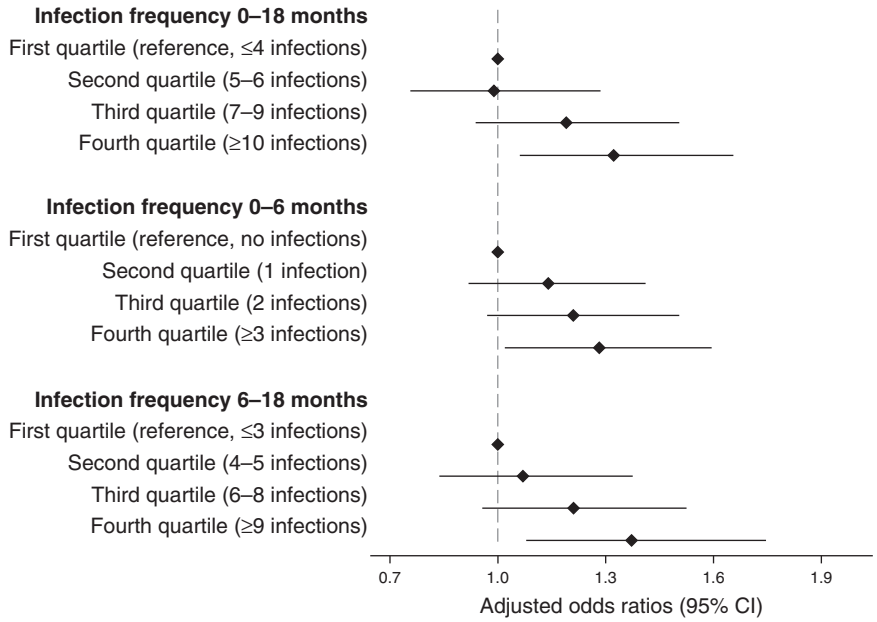

Figure 2. Risk of celiac disease according to infection frequency in the first 18 months of life. Odds ratios adjusted for attained age at the end of follow-up and sex.

infections up to age 18 months were significantly associated with later CD, gastrointestinal infections yielded similar, but nonsignificant, increased relative risk estimates.

Strengths and limitations

A major strength in this study is the use of prospectively recorded exposure and outcome data, minimizing the risk of recall bias. This study included >87,000 children that gave precise relative risk estimates and enabled us to perform important subanalyses, such as stratification per type of infection and age-period of infection exposure. Through the combined data from comprehensive questionnaires and national registries, we were able to adjust for several potential confounders, such as infant feeding practices and maternal CD.

We regard the risk of misclassification of diagnosed CD as low. In a recent validation of CD in the MoBa cohort, >92% of the CD diagnoses were confirmed. A high specificity for CD (i.e., few false-positive cases) is essential, which otherwise would result in attenuated relative risk estimates. Further, identification of CD via the combination of parental questionnaires and the nationwide NPR should give a high sensitivity for diagnosed CD. During a median follow-up period of 8.5 years (range, 4.5–14.5), the CD prevalence in this study was 0.8%, which is similar to previously reported prevalence figures of clinically diagnosed CD in children (4,29). Still, it is likely that some children (in particular among those with the shortest follow-up) will be diagnosed with CD after that age. Of greater concern is that the children in our study were not screened for CD. Consequently, some children may have an undiagnosed CD (i.e., be false-negative). Although the prevalence of undiagnosed CD should not exceed 1–2% (29), given the restriction of this study to children with clinically diagnosed CD, we cannot rule out that surveillance bias and reverse causation might somewhat have influenced our results.

We identified infectious disease data through parent reporting. It is unusual for Norwegian parents to seek medical advice when their child, including infants, experiences minor infections. Therefore, parent-reported infectious disease data should, as compared with inpatient records, better capture the average exposure to early life infections. Notably, the overall infection frequency in our study is almost identical to that reported from a Swedish cohort study based on parent-reported diary data collected prospectively in the child's first year of life (4). In addition, studies based on inpatient infectious disease data may be more susceptible to surveillance bias and inclusion of children with a high degree of comorbidity, leading to exaggerated relative risk estimates.

In MoBa, 40% of the invited women consented to participate. A comparison with all women giving birth in Norway at the time of study inclusion has shown that the cohort participants were older and less likely to smoke during pregnancy (30). However, these differences were not found to influence exposure-outcome associations between the MoBa-participants and the total population (30). Overall, we consider it unlikely that selection bias has influenced the internal validity of our study. Because of the self-selection in MoBa our results might be less generalizable to certain underrepresented populations—e.g., families with low socioeconomic position. This study was confined to children, and future studies are needed to elucidate whether prior infection frequency may be associated with adult CD.

Similar to other birth cohorts (31), MoBa suffers from drop-out. However, the consistency of our results between the analyses on children followed up to 6 and 18 months of age argues against selective drop-out, i.e., that infection frequency or, although rare, early CD manifestations have had an effect on the parents

Table 2. Risk of subsequent celiac disease according to infection frequency in the first 18 months of life

	Celiac disease (%)	No celiac disease (%)	Model I[a]		Model II[b]	
			aOR	95% CI	aOR	95% CI
Any infection[c]						
Per episode			1.03	1.02–1.05	1.03	1.01–1.05
0–4 episodes	133/581 (22.9)	18,282/72,330 (25.3)		Reference		Reference
5–6 episodes	103/581 (17.7)	14,621/72,330 (20.2)	0.99	0.76–1.28	0.96	0.74–1.25
7–9 episodes	156/581 (26.9)	18,604/72,330 (25.7)	1.19	0.94–1.50	1.15	0.91–1.46
10 or more episodes	189/581 (32.5)	20,823/72,330 (28.8)	1.32	1.06–1.65	1.27	1.01–1.60
Type of infection						
Gastroenteritis[d]						
Per episode			1.05	0.99–1.11	1.05	0.99–1.11
No	233/578 (40.3)	28,240/72,110 (39.2)		Reference		Reference
1 episode	168/578 (29.1)	22,729/72,110 (31.5)	0.90	0.74–1.10	0.88	0.71–1.08
2 or more episodes	177/578 (30.6)	21,141/72,110 (29.3)	1.04	0.86–1.27	1.03	0.84–1.26
Upper respiratory tract infection[e]						
Per episode			1.03	1.02–1.05	1.03	1.01–1.05
0–3 episodes	127/581 (21.9)	17,291/72,324 (23.9)		Reference		Reference
4–5 episodes	115/581 (19.8)	17,322/72,324 (24.0)	0.92	0.71–1.18	0.92	0.71–1.19
6–7 episodes	121/581 (20.8)	14,863/72,324 (20.6)	1.14	0.89–1.47	1.11	0.86–1.43
8 or more episodes	218/581 (37.5)	22,848/72,324 (31.6)	1.37	1.10–1.70	1.32	1.05–1.65
Lower respiratory tract infection[f]						
Per episode			1.12	1.01–1.23	1.10	0.99–1.22
No	475/579 (82.0)	61,355/72,145 (85.0)		Reference		Reference
At least one episode	104/579 (18.0)	10,790/72,145 (15.0)	1.29	1.04–1.59	1.27	1.02–1.59

aOR, adjusted odds ratio; 95% CI, 95% confidence interval.

[a]Odds ratios estimated through logistic regression adjusted for child's age at the end of follow-up and sex (model I). Because of internal attrition the number of children varied between analyses (children with celiac disease: 581–578; children without celiac disease: 72,330–71,980).

[b]In model II, we also adjusted for maternal smoking in pregnancy, education level, maternal celiac disease, parity, birth weight, prematurity, age at the time of gluten introduction, and duration of breastfeeding. Children with complete data on covariates were included in model II. Because of internal attrition the number of children varied between analyses (children with celiac disease: 554–551; children without celiac disease: 68,503–68,191).

[c]Includes parent-reported gastroenteritis, upper or lower respiratory tract infection, and urinary tract infection. The variable was divided into quartiles.

[d]The variable was divided into tertiles.

[e]Includes common cold, throat infection, ear infection, and croup. The variable was divided into quartiles.

[f]Defined as pneumonia or bronchitis.

questionnaire completion. The timing of infections was defined in 6-month intervals, which limited our ability to examine infections occurring at the time of gluten introduction. Finally, in a nationwide study, some of the individual data may be missing or recorded erroneously, and we were unable to verify the parent-reported infectious disease data against medical charts. However, given the prospective data collection, potential misclassification of data should not differ by future CD status and therefore not cause spurious association but only bias our results toward the null.

Interpretation of findings and previous literature

Most (6–10), but not all (4,5), previous studies have reported significant positive associations between early life infection and later CD, with relative risk estimates ranging from 1.46 (10) to 1.94 (6). These studies have mostly been restricted to inpatient infectious disease (5,8–10), and thereby more susceptible to surveillance bias that may inflate risk estimates. Our results are consistent with the previously reported positive associations, but we contribute by showing that this association is not limited to inpatient infectious disease and by presenting risk estimates per type of infection. This study also demonstrates that the association between infectious disease and later CD is not due to potential confounders that previous studies were unable to adjust for, e.g., breastfeeding.

Our results are in line with existing immunological data suggesting that repeated infections may trigger CD (3). Some of the hypotheses for an infectious etiology of CD include a virus-medi-

Table 3. Risk of subsequent celiac disease according to doctors visits and/or hospital admission for infection in the first 18 months of life

	Celiac disease (%)	No celiac disease (%)	Model i[a]		Model II[b]	
			aOR	95% CI	aOR	95% CI
Any infection[c]						
No	364/580 (62.8)	48,615/72,322 (67.2)		Reference		Reference
At least one episode	216/580 (37.2)	23,707/72,322 (32.8)	1.25	1.06–1.48	1.25	1.05–1.49
No	364/580 (62.8)	48,615/72,322 (67.2)		Reference		Reference
1 episode	156/580 (26.9)	18,009/72,322 (24.9)	1.19	0.98–1.43	1.20	0.99–1.45
2 or more episodes	60/580 (10.3)	5698/72,322 (7.9)	1.46	1.10–1.92	1.42	1.06–1.90
Type of infection						
Gastroenteritis						
No	549/576 (95.3)	68,807/71,860 (95.8)		Reference		Reference
At least one episode	27/576 (4.7)	3053/71,860 (4.2)	1.13	0.77–1.67	1.15	0.77–1,70
Upper respiratory tract infection[d]						
No	401/580 (69.1)	52,977/72,313 (73.3)		Reference		Reference
At least one episode	179/580 (30.9)	19,336/72,313 (26.7)	1.25	1.05–1.50	1.25	1.04–1.50
Lower respiratory tract infection[e]						
No	528/579 (91.2)	67,505/72,107 (93.6)		Reference		Reference
At least one episode	51/579 (8.8)	4602/72,107 (6.4)	1.48	1.10–1.97	1.40	1.03–1.92

aOR, adjusted odds ratio; 95% CI, 95% confidence interval.

[a]Odds ratios estimated through logistic regression adjusted for child's age at the end of follow-up and sex (model I). Because of internal attrition the number of children varied between analyses (children with celiac disease: 580–576; children without celiac disease: 72,322–71,860).

[b]In model II, we also adjusted for maternal smoking in pregnancy, education level, maternal celiac disease, parity, birth weight, prematurity, age at the time of gluten introduction, and duration of breastfeeding. Children with complete data on covariates were included in model II. Because of internal attrition the number of children varied between analyses (children with celiac disease: 553–549; children without celiac disease: 68,494–68,086).

[c]Doctors visits and/or hospital admission for gastroenteritis, upper or lower respiratory tract infection, and urinary tract infection.

[d]Doctors visits and/or hospital admission for common cold, throat infection, ear infection, or croup.

[e]Doctors visits and/or hospital admission for pneumonia or bronchitis.

ated induction of pro-inflammatory type-1 interferons (11), or virus-induced upregulation and release of the enzyme tissue transglutaminase (12). Further, infections involving the intestinal tract may increase intestinal permeability and thereby facilitate epithelial translocation of gluten, a key process early in CD pathogenesis (13). Our finding of a significant association with respiratory tract infection is consistent with a previous study reporting that children with CD were more likely to have attended hospital for a prior viral bronchiolitis (10). In children, respiratory viruses frequently cause gastrointestinal manifestations (32) and may thereby also contribute in inducing the intestinal processes outlined above hypothesized to trigger CD. Still, we urge caution when interpreting the differences in ORs per specific infectious type, as the CIs overlapped and the type of infection of the child may have been misconstrued by the parent.

In 2006, Stene *et al.* found that children with a raised frequency of serologically defined rotavirus infections had a significantly increased risk of CD autoimmunity (6). We did not find a significant association between symptomatic gastroenteritis and later CD diagnosis. However, we remain cautious in the interpretation of this nonsignificant association. First, the current study lacked data on specific pathogens, and, we were unable to study subclinical infections, which are common in infancy (33). Thus, we cannot rule out that the inclusion of gastrointestinal infections occurring without notably symptoms may have yielded disparate relative risk estimates nor the possibility that gastroenteritis caused by rotavirus, or other specific pathogens, constitutes an independent risk factor for CD. Second, gastroenteritis is relatively infrequent up to 6–12 months of age (34), and the lack of significance in our study may be attributable to insufficient statistical power for this infectious type. On the other hand, our nonsignificant association is consistent with previous data (4,5,7), arguing against gastroenteritis as a major risk factor for CD.

Another explanation for our findings may be genetic or environmental factors acting as common causes of CD and infection susceptibility. Several studies have found that celiac patients are more likely to contract infectious diseases (35), in particular respiratory tract infections (36,37), as compared with the general population. Although this increased infection susceptibility has been explained by impaired nutritional status or functionally hyposplenism in CD patients (38,39), an alternative explanation may be a CD-associated genetic susceptibility for infectious

disease. The human leukocyte antigen genotype, crucial for the innate immune system's recognition of self and non-self, is the prototypical candidate gene of infection susceptibility and the single most important predictor of CD (40,41), However, to our knowledge, there are no consistent data supporting a relationship between CD-associated high-risk human leukocyte antigen genotypes and increased susceptibility to common childhood infections. Conversely, it has been hypothesized that autoimmune-prone human leukocyte antigen haplotypes might have been positive selected in human evolution based on their ability to clear infections (42). Further, unpublished data from our research group indicate that CD-risk haplotype human leukocyte antigen-DQ2.5 (predicted using tag single nucleotide polymorphism rs2187668 (43) in 506 MoBa-children without CD) is not significantly associated with parent-reported infection frequency up to age 18 months (P=0.56).

Recently, neonatal vitamin D deficiency was associated with later viral bronchiolitis, and vitamin D deficiency is overrepresented in individuals with untreated CD (44,45). To our knowledge, there are no studies on a causal role of vitamin D in CD and the fact that adjustment for birth season did not influence our results does not support vitamin D status as a confounder.

We cannot rule out that children with high infection frequency or infections requiring medical care may be more likely to undergo investigation for CD, leading to surveillance bias. However, the consistency of our results between children with infections before age 6 months (when CD is virtually non-existent) and after age 6 months, as well as the nonsignificant association for any gastroenteritis nor gastroenteritis requiring medical care, argues against surveillance bias as a sole cause of our results.

In conclusion, we found a modest but significantly increased risk for later CD according to infection frequency in the first 18 months of life. Although respiratory tract infections up to age 18 months were significantly associated with later CD, gastrointestinal infections yielded similar, but nonsignificant, increased relative risk estimates. Our results complement and extend those from previous epidemiological studies and are in line with immunological data suggesting that early life infections may have a role in CD development. However, non-causal explanations for this association due to surveillance bias and reverse causation cannot be excluded.

ACKNOWLEDGMENTS

We are grateful to all the participating families in Norway who took part in this on-going cohort study.

CONFLICT OF INTEREST

Guarantor of the article: Karl Mårild, MD, PhD.

Specific author contributions: K.M.; C.R.K.; G.T.; L.C.S.; K.S. read and met the ICMJE criteria for authorship. K.M.; C.R.K.; G.T.; L.C.S.; K.S. agree with the results and conclusions of the manuscript. K.M.; K.S. designed the study. KS collected the data. K.M. analyzed the data and wrote the first draft of the paper. C.R.K.; G.T.; L.C.S.; K.S. contributed to the writing of the paper. C.R.K.; G.T.; L.C.S. contributed to the design of the study and interpretation of the data analyses. K.M.; C.R.K.; G.T.; L.C.S.; K.S. interpreted the data; approved the final version of the manuscript. K.M. was responsible for data integrity. K.S.; L.C.S. supervised the project including data analyses. L.C.S.; K.S. obtained funding.

Financial support: The Norwegian Mother and Child Cohort Study (MoBa): the Norwegian Ministry of Health and the Ministry of Education and Research, NIH/NIEHS (contract no. N01-ES-75558), NIH/NINDS (grant no.1 UO1 NS 047537-01 and grant no.2 UO1 NS 047537-06A1), and the Norwegian Research Council/FUGE (grant no. 151918/S10). This subproject within MoBa is supported by the Norwegian Research Council (grant no. 221909/F20). K.S.: Oak Foundation; CRK: Østfold Hospital Trust. The funding sources did not influence the study's design, conduct, and reporting.

Potential competing interests: None.

Study Highlights

WHAT IS CURRENT KNOWLEDGE

- ✓ Celiac disease (CD) is an immune-mediated disease prevalent in 1–2% of Caucasians.
- ✓ Although there are strong indications that environmental factors influence CD development, the nature and timing of such factors are largely unknown.
- ✓ Studies on early life infections and risk of CD are inconsistent but have mostly been limited to retrospective designs, inpatient data, or insufficient statistical power.

WHAT IS NEW HERE

- ✓ Children with high infection frequency in the first 18 months of life had a significantly increased risk of later CD, with an increasing risk according to the number of infections.
- ✓ This is the first large-scale population-based cohort study of this association.
- ✓ We present odds ratios for later CD according to type and severity of infections acquired before age 6 months or at age 6–18 months.
- ✓ In contrast to previous studies in this field, we demonstrate that the association between early life infections and CD remained largely unaffected by the confounding influences of antibiotic treatment or breastfeeding duration, etc.

REFERENCES

1. Walker MM, Murray JA, Ronkainen J *et al.* Detection of celiac disease and lymphocytic enteropathy by parallel serology and histopathology in a population-based study. Gastroenterology 2010;139:112–9.
2. Ludvigsson JF, Montgomery SM, Ekbom A *et al.* Small-intestinal histopathology and mortality risk in celiac disease. JAMA 2009;302:1171–8.
3. Abadie V, Sollid LM, Barreiro LB *et al.* Integration of genetic and immunological insights into a model of celiac disease pathogenesis. Annu Rev Immunol 2011;29:493–525.
4. Welander A, Tjernberg AR, Montgomery SM *et al.* Infectious disease and risk of later celiac disease in childhood. Pediatrics 2010;125:e530–6.
5. Mårild K, Stephansson O, Montgomery S *et al.* Pregnancy outcome and risk of celiac disease in offspring: a nationwide case-control study. Gastroenterology 2012;142:39–45.
6. Stene LC, Honeyman MC, Hoffenberg EJ *et al.* Rotavirus infection frequency and risk of celiac disease autoimmunity in early childhood: a longitudinal study. Am J Gastroenterol 2006;101:2333–40.

7. Myleus A, Hernell O, Gothefors L *et al.* Early infections are associated with increased risk for celiac disease: an incident case-referent study. BMC Pediatr 2012;12:194.
8. Sandberg-Bennich S, Dahlquist G, Källen B. Coeliac disease is associated with intrauterine growth and neonatal infections. Acta Paediatr 2002;91:30–3.
9. Canova C, Zabeo V, Pitter G *et al.* Association of maternal education, early infections, and antibiotic use with celiac disease: a population-based birth cohort study in northeastern Italy. Am J Epidemiol 2014;180:76–85.
10. Tjernberg AR, Ludvigsson JF. Children with celiac disease are more likely to have attended hospital for prior respiratory syncytial virus infection. Dig Dis Sci 2014;59:1502–8.
11. Sollid LM, Jabri B. Triggers and drivers of autoimmunity: lessons from coeliac disease. Nat Rev Immunol 2013;13:294–302.
12. Jabri B, Sollid LM. Tissue-mediated control of immunopathology in coeliac disease. Nat Rev Immunol 2009;9:858–70.
13. Fasano A, Shea-Donohue T. Mechanisms of disease: the role of intestinal barrier function in the pathogenesis of gastrointestinal autoimmune diseases. Nat Clin Pract Gastroenterol Hepatol 2005;2:416–22.
14. Magnus P, Irgens LM, Haug K *et al.* Cohort profile: The Norwegian Mother and Child Cohort Study (MoBa). Int J Epidemiol 2006;35:1146–50.
15. Størdal K, Bakken IJ, Suren P *et al.* Epidemiology of coeliac disease and comorbidity in Norwegian children. J Pediatr Gastroenterol Nutr 2013;57:467–71.
16. Emilsson L, Magnus MC, Størdal K. Perinatal risk factors for development of celiac disease in children, based on The Prospective Norwegian Mother and Child Cohort Study. Clin Gastroenterol Hepatol 2015;13:921–7.
17. Paricio Talayero JM, Lizan-Garcia M, Otero Puime A *et al.* Full breast-feeding and hospitalization as a result of infections in the first year of life. Pediatrics 2006;118:e92–9.
18. Ludvigsson JF, Ludvigsson J. Parental smoking and risk of coeliac disease in offspring. Scand J Gastroenterol 2005;40:336–42.
19. Kramer MS, Chalmers B, Hodnett ED *et al.* Promotion of Breastfeeding Intervention Trial (PROBIT): a randomized trial in the Republic of Belarus. JAMA 2001;285:413–20.
20. Størdal K, White RA, Eggesbo M. Early feeding and risk of celiac disease in a prospective birth cohort. Pediatrics 2013;132:e1202–9.
21. Ivarsson A, Hernell O, Stenlund H *et al.* Breast-feeding protects against celiac disease. Am J Clin Nutr 2002;75:914–21.
22. Mårild K, Ye W, Lebwohl B *et al.* Antibiotic exposure and the development of coeliac disease: a nationwide case-control study. BMC Gastroenterol 2013;13:109.
23. Irgens LM. The Medical Birth Registry of Norway. Epidemiological research and surveillance throughout 30 years. Acta Obstet Gynecol Scand 2000;79:435–9.
24. Lebwohl B, Green PH, Murray JA *et al.* Season of birth in a nationwide cohort of coeliac disease patients. Arch Dis Child 2013;98:48–51.
25. Ivarsson A, Hernell O, Nyström L *et al.* Children born in the summer have increased risk for coeliac disease. J Epidemiol Community Health 2003;57:36–9.
26. Wingren CJ, Agardh D, Merlo J. Revisiting the risk of celiac disease in children born small for gestational age: a sibling design perspective. Scand J Gastroenterol 2012;47:632–9.
27. Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. Biometrics 1986;42:121–30.
28. Norwegian Insititute of Public Health. Childhood Immunisation Programme in Norway. 2014 [cited 25 June 2015]; Available from http://www.fhi.no/artikler/?id=90915.
29. Myleus A, Ivarsson A, Webb C *et al.* Celiac disease revealed in 3% of Swedish 12-year-olds born during an epidemic. J Pediatr Gastroenterol Nutr 2009;49:170–6.
30. Nilsen RM, Vollset SE, Gjessing HK *et al.* Self-selection and bias in a large prospective pregnancy cohort in Norway. Paediatr Perinat Epidemiol 2009;23:597–608.
31. Olsen J, Melbye M, Olsen SF *et al.* The Danish National Birth Cohort–its background, structure and aim. Scand J Public Health 2001;29:300–7.
32. Clark B, McKendrick M. A review of viral gastroenteritis. Curr Opin Infect Dis 2004;17:461–9.
33. Phillips G, Lopman B, Rodrigues LC *et al.* Asymptomatic rotavirus infections in England: prevalence, characteristics, and risk factors. Am J Epidemiol 2010;171:1023–30.
34. Giaquinto C, van Damme P. Age distribution of paediatric rotavirus gastroenteritis cases in Europe: the REVEAL study. Scand J Infect Dis 2010;42:142–7.
35. Ludvigsson JF, Olen O, Bell M *et al.* Coeliac disease and risk of sepsis. Gut 2008;57:1074–80.
36. Mårild K, Fredlund H, Ludvigsson JF. Increased risk of hospital admission for influenza in patients with celiac disease: a nationwide cohort study in Sweden. Am J Gastroenterol 2010;105:2465–73.
37. Ludvigsson JF, Sanders DS, Maeurer M *et al.* Risk of tuberculosis in a large sample of patients with coeliac disease--a nationwide cohort study. Aliment Pharmacol Ther 2011;33:689–96.
38. Di Sabatino A, Rosado MM, Cazzola P *et al.* Splenic hypofunction and the spectrum of autoimmune and malignant complications in celiac disease. Clin Gastroenterol Hepatol 2006;4:179–86.
39. Malandrino N, Capristo E, Farnetti S *et al.* Metabolic and nutritional features in adult celiac patients. Dig Dis 2008;26:128–33.
40. Lionetti E, Castellaneta S, Francavilla R *et al.* Introduction of gluten, HLA status, and the risk of celiac disease in children. N Engl J Med 2014;371:1295–303.
41. Blackwell JM, Jamieson SE, Burgner D. HLA and infectious diseases. Clin Microbiol Rev 2009;22:370–85.
42. Mangalam AK, Taneja V, David CS. HLA class II molecules influence susceptibility vs. protection in inflammatory diseases by determining the cytokine profile. J Immunol 2013;190:513–8.
43. Monsuur AJ, de Bakker PI, Zhernakova A *et al.* Effective detection of human leukocyte antigen risk alleles in celiac disease using tag single nucleotide polymorphisms. PLoS One 2008;3:e2270.
44. Margoni D, Chouliaras G, Duscas G *et al.* Bone health in children with celiac disease assessed by dual x-ray absorptiometry: effect of gluten-free diet and predictive value of serum biochemical indices. J Pediatr Gastroenterol Nutr 2012;54:680–4.
45. Belderbos ME, Houben ML, Wilbrink B *et al.* Cord blood vitamin D deficiency is associated with respiratory syncytial virus bronchiolitis. Pediatrics 2011;127:e1513–20.