



ORIGINAL ARTICLE

Geographical variation and incidence of inflammatory bowel disease among US women

Hamed Khalili,¹ Edward S Huang,¹ Ashwin N Ananthakrishnan,¹ Leslie Higuchi,² James M Richter,¹ Charles S Fuchs,^{3,4} Andrew T Chan^{1,4}

¹Division of Gastroenterology, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, USA

²Division of Gastroenterology, Children's Hospital, Boston, Massachusetts, USA

³Department of Adult Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts, USA

⁴Channing Laboratory, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA

Correspondence to

Dr Andrew T Chan, Division of Gastroenterology, Gastrointestinal Unit, Massachusetts General Hospital, 55 Fruit Street, GRJ-728A, Boston, MA 02114, USA; achan@partners.org

Revised 24 November 2011
Accepted 30 November 2011

ABSTRACT

Objective Geographical variation in the incidence of Crohn's disease (CD) and ulcerative colitis (UC) according to the latitude of residence has been reported in Europe. However, there are no comparable data in the USA. The incidence of CD and UC in relation to latitude was assessed in a geographically diverse population of women enrolled in two large prospective studies in the USA.

Design A prospective study was undertaken of women enrolled in the Nurses' Health Study I (NHS) in 1976 and in the NHS II in 1989. Information on state of residence at the time of birth, at age 15 years and age 30 years was collected in 1992 in NHS I and in 1993 in NHS II. Reported diagnoses of incident CD or UC to the end of 2003 were confirmed by medical record review. Cox proportional hazards models were used to calculate HRs and 95% CIs for risk of CD and UC.

Results In both cohorts, among 175 912 women reporting their residence in 1992, 257 cases of CD and 313 cases of UC were documented over 3 428 376 person-years of follow-up. The incidence of CD and UC increased significantly with increasing latitude ($p_{\text{trend}} < 0.01$), with residence at age 30 years more strongly associated with risk. Compared with women residing in northern latitudes at age 30, the multivariate-adjusted HR for women residing in southern latitudes was 0.48 (95% CI 0.30 to 0.77) for CD and 0.62 (95% CI 0.42 to 0.90) for UC. The effect of latitude of residence on risk of CD and UC did not vary according to smoking history ($p_{\text{interaction}} = 0.26$ for CD and 0.99 for UC).

Conclusion In a population of US women, increasing latitude of residence was associated with a higher incidence of CD and UC.

INTRODUCTION

Crohn's disease (CD) and ulcerative colitis (UC), collectively known as inflammatory bowel disease (IBD), are archetypical inflammatory diseases in which a barrier normally maintained by adaptive and innate immunity is disrupted. Despite recent advances in genetics and identification of more than 100 risk loci for UC and CD,^{1,2} the pathogenesis of the two diseases remains largely unknown. Because the overall population risk attributable to known genetic risk loci is low, there are likely environmental and lifestyle factors that are important in disease pathogenesis.

Geographical variation in the incidence of CD or UC may reflect variation in environmental risk factors for the disease. Previous studies in Europe

Significance of this study

What is already known about the topic?

- ▶ Geographical variation according to latitude of residence in the incidence of Crohn's disease (CD) and ulcerative colitis (UC) has been reported in Europe.
- ▶ It is unclear if this geographical variation in European populations would also be evident in a more genetically diverse population such as the USA.
- ▶ Previous studies did not have the opportunity to examine latitude in relation to risk of CD or UC using time points with more plausible biological latency and have relied on location at the time of diagnosis.

What are the new findings?

- ▶ In two large prospective cohorts of US women, the incidence of UC and CD was significantly lower among women who resided in the southern latitudes than in those residing in the northern latitudes.
- ▶ Residence later in life (age 30 years) was more strongly related to risk.

How might it impact on clinical practice in the foreseeable future?

- ▶ Our results support the importance of biological pathways that mediate these geographical differences in the pathogenesis of inflammatory bowel disease.
- ▶ Understanding such pathways could eventually lead to the development of novel lines of therapy as well as interventions that may modulate the risk of incident disease.

have linked latitude and geographical variations to risk of CD and UC.³⁻⁶ However, these studies are limited by consideration of geographical location only at time of diagnosis, inability to concurrently examine other lifestyle factors and lack of physician confirmation of UC and CD diagnoses. Moreover, it is unclear if geographical variation in European populations would also be evident in a more genetically diverse population such as the USA. The only large study of geographical variation on risk of UC and CD in the USA was limited to a Medicare population with only 2 years of follow-up.⁷

We therefore sought to examine the association between geographical variation according to

latitude of residence and risk of CD and UC in two large ongoing prospective studies of US women, the Nurses' Health Study I (NHS I) and Nurses' Health Study II (NHS II). With more than 20 years of biennially updated data on lifestyle, these prospective cohorts offered us the unique opportunity to examine the association between latitude of residence at several different time points in early life and subsequent risk of UC and CD in the context of other risk factors.

METHODS

Study population

The NHS I is a prospective cohort that began in 1976 when 121 700 US female registered nurses aged 30–55 years completed a mailed health questionnaire. Follow-up questionnaires are mailed every 2 years to update health information. In 1989 a parallel cohort, the NHS II, enrolled 116 686 US female nurses aged 25–42 years. These women have been followed with similar biennial questionnaires. Follow-up in both cohorts has consistently exceeded 90%.

Assessment of geographical location

Participants reported their state of residence at birth, at age 15 years and at age 30 years in the 1992 questionnaire in the NHS I and similarly in the 1993 questionnaire in the NHS II; 73% of all women in NHS I and 78% of all women in NHS II provided information about their states of residence at all three time points. Consistent with previous studies,^{8,9} we divided the continental USA into northern, middle and southern tiers of latitude. The northern tier includes states north of 41–42° latitude in the east (CT, ME, MA, NH, NY, RI, VT), central (MI, MN, WI), mountain (ID, MT, NE, ND, SD, WY) and Pacific (OR, WA) time zones. The southern tier consists of states lying south of 37° latitude in the east (FL, GA, NC, SC), central (AL, AR, LA, MS, TN), mountain (AZ, NM, OK, TX) and Pacific (southern CA) zones. Finally, the middle tier consists of the remaining states in the east (DE, DC, MD, NJ, PA, VA, WV), central (IL, IN, IA, KY, MO, OH), mountain (CO, KS, NV, UT) and Pacific (northern CA) zones. Hawaii and Puerto Rico were included in the southern tier while Alaska was included in the northern tier.

Outcome ascertainment

Since the initial baseline questionnaires in 1976 in NHS I and in 1989 in NHS II, participants have reported diagnoses of UC or CD through an open-ended response on biennial surveys. In NHS I we have specifically queried participants about diagnoses of UC since 1982 and CD since 1992. In NHS II we have specifically queried participants about diagnoses of both UC and CD since 1993. In both cohorts, when a diagnosis was reported on any biennial questionnaire, related medical records were requested and reviewed by two gastroenterologists blinded to exposure information. Data were extracted on diagnostic tests, histopathology, anatomical location of disease and disease behaviour. Using standardised criteria, patients were considered to have UC based on a typical clinical presentation for ≥ 4 weeks and endoscopic or surgical pathological specimen consistent with UC (eg, evidence of chronicity). A diagnosis of CD was made based on typical clinical history for ≥ 4 weeks along with at least one characteristic change on endoscopic, surgical or radiological examination demonstrating small bowel involvement, or an endoscopic or surgical pathological specimen consistent with CD (eg, transmural inflammation or granulomatous disease). Disagreements were resolved through consensus.

For this study, 1545 women reported a diagnosis of UC or CD on at least one biennial questionnaire from June 1976 to June 2002 in NHS I and 1256 women reported a diagnosis of UC or CD on at least one biennial questionnaire from June 1989 to June 2003 in NHS II. From these women, 1310 (85%) in NHS I and 1125 (90%) in NHS II responded to a supplementary questionnaire. After excluding subjects who subsequently denied the diagnosis of UC or CD on the supplementary questionnaire (n=571 in NHS I; n=475 in NHS II) or denied permission for record review (n=191 in NHS I; n=182 in NHS II), we requested medical records from 548 women in NHS I and 468 women in NHS II. Among those from whom we requested medical records, we obtained 481 (88%) records in NHS I and 436 (93%) records in NHS II with adequate information for review (68 cases in NHS I and 32 cases in NHS II with inadequate information to confirm diagnoses). We confirmed incident cases of UC, CD and chronic colitis (indeterminate colitis, microscopic colitis) in 373 NHS I participants and 324 NHS II participants for a case confirmation rate of 78% in NHS I and 74% in NHS II of the medical records reviewed.

For all analyses we excluded individuals with unknown date of diagnosis (n=2) and cases of indeterminate or microscopic colitis (n=31), leaving a total of 284 cases of CD and 332 cases of UC to the end of follow-up (1 June 2002 for NHS I and 1 June 2003 for NHS II). For our analysis of latitude and risk of UC and CD, we excluded participants who did not return the 1992 questionnaire in NHS I or 1993 questionnaire in NHS II which queried participants about state of residence (n=23 in NHS I and n=19 in NHS II), leaving 257 incident cases of CD and 313 incident cases of UC to the end of follow-up.

The baseline characteristics of participants for whom we were able to obtain complete medical records were similar to those of participants for whom we were unable to obtain sufficient records (mean age, 57.1 vs 57.1 years; non-white, 5% vs 5%; current smoker, 22.2% vs 19.4%; mean body mass index, 25.1 kg/m² vs 24.9 kg/m²; current or past oral contraceptive use, 53.7% vs 55.8%; postmenopausal hormone use, 42.7% vs 44.9%; southern latitude, 15.1% vs 14.9%, $p > 0.15$ for all comparisons). Women for whom we did not confirm CD or UC were included in the analyses as non-cases.

Other covariates

On each biennial questionnaire, women were asked about several lifestyle factors including body weight, smoking status, menopause status, hormonal replacement therapy and oral contraceptive use. Participants' self-report of body weight has been previously validated.¹⁰ At the initiation of the NHS I in 1976 and NHS II in 1989, women were enrolled without regard to race or ethnicity. In the 1992 questionnaire in NHS I and baseline questionnaire in NHS II, participants were asked whether they had the following ancestries: African, Asian, Hispanic, Scandinavian, southern European/Mediterranean, other white or other ancestry. Most women reported white ancestries (93% in NHS I and 91% in NHS II), reflecting the under-representation of minorities in the nursing profession at the inception of both cohorts. More than 90% of the women reported a single ancestry. We categorised participants as southern European/Mediterranean or as Scandinavian when that was the only ancestry reported, as other white when a mixture of only white ancestries was reported, and as non-white when African, Asian or Hispanic ancestry was reported. More than 85% of women in NHS I and 99% of women in NHS II provided information about their ancestry.

Statistical analysis

After excluding women with a history of CD or UC prior to the baseline questionnaires, we calculated overall incidence rates for CD and UC according to 5-year categories of age between 1976 and 2002 for NHS I and between 1989 and 2003 for NHS II. To examine the risk of CD or UC according to geographical residence, we limited our analysis to the 175 912 women (86 710 in NHS I and 89 202 in NHS II) from whom we had complete information on state of residence at birth, age 15 and age 30. We used a χ^2 test for categorical variables and Wilcoxon rank sum test for continuous variables to compare baseline characteristics according to the latitude of residence. Person-time for each participant was calculated from the date of return of their baseline questionnaires to the date of the diagnosis of UC or CD, date of last questionnaire, death from any cause or 1 June 2002 for NHS I and 1 June 2003 for NHS II, whichever came first. Cox proportional hazards modelling with time-varying covariates was used to adjust for other known or suspected risk factors prior to each 2-year interval to calculate adjusted HRs and 95% CI. Because we observed no heterogeneity in the association of latitude with CD or UC in separate analyses of NHS I and NHS II (p for heterogeneity >0.30 for both UC and CD), we combined data from the two cohorts. We also examined the association between latitude and risk of UC or CD according to strata of smoking and evaluated for potential interaction using cross-classified categories of smoking and latitude. We tested the significance of this interaction by using the log likelihood ratio test comparing the model with these cross-classified categories with a model that included smoking as an independent variable. SAS V.9.1.3 was used for these analyses. All p values were two-sided and <0.05 was considered statistically significant.

RESULTS

Incidence of UC and CD in NHS I and II

Among the 103 691 women in NHS I and 114 971 women in NHS II cohorts without a history of CD or UC prior to enrolment, we confirmed 284 cases of CD and 332 cases of UC over 4 209 454 years of follow-up. The age-adjusted incidence of CD per 100 000 person-years ranged from 5.6 to 10.3 and the age-adjusted incidence of UC per 100 000 person-years ranged from 6.2 to 12.9 (table 1). There appeared to be a bimodal pattern of incidence, such that the highest incidence rates for both CD and UC were in women aged <30 years or >60 years of age. These incidence rates are largely consistent with those reported by other US population-based cohorts (figures 1 and 2).^{11 12}

Latitude of residence and risk of UC and CD

The baseline characteristics of 175 912 women who provided information about geographical residence are shown in table 2

Table 1 Age-specific incidence of Crohn's disease and ulcerative colitis in Nurses' Health Study (NHS) I and II cohorts

| Age group | Crohn's disease | | Ulcerative colitis | |
|-----------|-------------------|---------------------------------|--------------------|---------------------------------|
| | Cases/person-year | Incidence*/100 000 person-years | Cases/person-year | Incidence*/100 000 person-years |
| <30 | 7/70 005 | 10.0 | 9/70 005 | 12.9 |
| 30-39 | 52/881 741 | 5.9 | 80/881 741 | 9.1 |
| 40-49 | 78/1 398 110 | 5.6 | 96/1 398 110 | 6.9 |
| 50-59 | 65/1 047 118 | 6.2 | 77/1 047 118 | 7.4 |
| 60-69 | 64/618 440 | 10.3 | 58/618 440 | 9.4 |
| ≥70 | 18/194 040 | 9.3 | 12/194 040 | 6.2 |

*Incidence from 1976 to 2002 in NHS I and from 1989 to 2003 in NHS II.

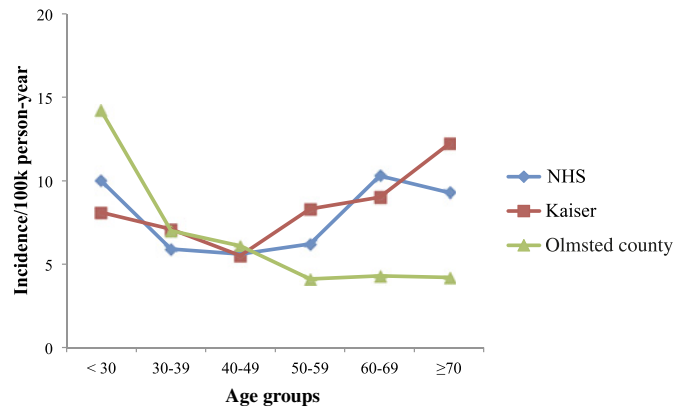


Figure 1 Comparison of age-specific incidence rates of Crohn's disease in the Nurses' Health Study (NHS), Kaiser Permanente and Olmsted County cohorts.^{11 12}

according to the latitude of residence at age 30. Compared with northern latitudes, women who resided in southern and middle latitudes at age 30 were more likely to have never smoked and were less likely to be current users of oral contraceptives.

Compared with women who resided in northern latitudes, women in southern latitudes had a lower age-adjusted risk of developing UC, which was not appreciably altered even after adjusting for ancestry as well as other known or suspected risk factors (table 3). The lower risk of UC associated with decreasing latitude appeared stronger according to residence at older ages. The multivariate-adjusted HRs for UC were 0.67 (95% CI 0.44 to 1.01) for women who resided in southern latitudes at age 15 and 0.62 (95% CI 0.42 to 0.90) for women who resided in southern latitudes at age 30. This effect was consistent, although somewhat attenuated according to latitude of residence at birth or age 15. Compared with women born in northern latitudes, the multivariate HRs associated with residence at birth for UC was 1.00 (95% CI 0.78 to 1.27) for women born in middle latitudes and 0.69 (95% CI 0.44 to 1.07) for women born in southern latitudes. This association was not materially changed when we restricted the analysis to women who resided in the same latitude at birth, age 15 and age 30. Compared with women who had lived consistently in northern latitudes, the multivariate-adjusted HR was 0.63 (95% CI 0.37 to 1.08) for women who lived consistently in southern latitudes.

We also evaluated the risk of CD according to latitude of residence (table 4). As with UC, the association between latitude

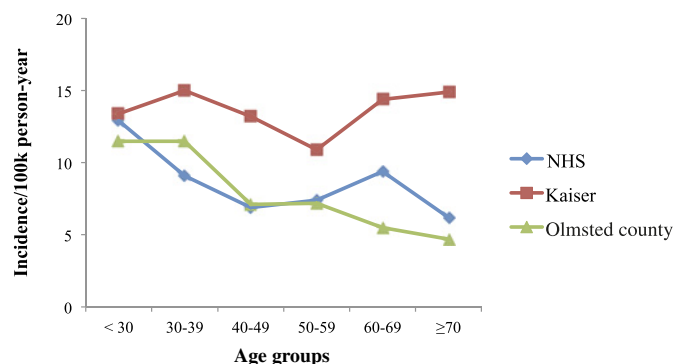


Figure 2 Comparison of age-specific incidence rates of ulcerative colitis in the Nurses' Health Study (NHS), Kaiser Permanente and Olmsted County cohorts.^{11 12}

Table 2 Baseline characteristics of women in the Nurses' Health Study (NHS) I and II by latitude of residence at age 30*

| | Northern latitudes (n = 61 923) | Middle latitudes (n = 84 286) | Southern latitudes (n = 29 703) |
|----------------------------------------------|---------------------------------|-------------------------------|---------------------------------|
| Mean age, years | 39.3 (7.2) | 38.9 (7.0) | 38.3 (6.9) |
| Ancestry/race, % | | | |
| Southern European/Mediterranean | 19.0 | 17.5 | 15.9 |
| Scandinavian | 7.3 | 6.5 | 9.3 |
| Other Caucasian | 66.9 | 69.4 | 65.0 |
| Non-white/unknown | 6.8 | 6.6 | 9.9 |
| Mean (SD) body mass index, kg/m ² | 23.8 (4.4) | 24.0 (4.7) | 23.6 (4.6) |
| Post-menopause, % | 16.4 | 16.3 | 17.6 |
| Hormonal replacement therapy, † % | | | |
| Never | 54.3 | 52.8 | 42.6 |
| Past | 18.3 | 17.2 | 17.6 |
| Current | 27.4 | 30.0 | 39.8 |
| Oral contraceptive use, % | | | |
| Never | 33.3 | 31.4 | 27.6 |
| Past | 54.3 | 57.6 | 62.5 |
| Current | 12.4 | 11.0 | 9.9 |
| Smoking, % | | | |
| Never | 49.6 | 57.4 | 60.6 |
| Past | 25.7 | 21.8 | 21.9 |
| Current | 24.7 | 20.9 | 17.5 |

*According to the baseline questionnaire for NHS I (1976) and NHS II (1989), except for ancestry and location of residence at age 30 which were derived from the 1992 questionnaire in NHS I and 1993 questionnaire in NHS II.
 †Among post-menopausal women.

and risk of CD was evident accounting for latitude of residence at age 30. Compared with women who resided in northern latitudes at age 30, the multivariate-adjusted HRs were 0.84

(95% CI 0.64 to 1.10) for women who resided in middle latitudes at age 30 and 0.48 (95% CI 0.30 to 0.77) for women who resided in southern latitudes at age 30. Although statistical power was limited, these effect estimates were not materially altered when we restricted the analyses to women who resided in the same latitudes at birth, age 15 and age 30. Compared with women who resided consistently in northern latitudes, the multivariate-adjusted HR was 0.77 (95% CI 0.58 to 1.04) for women who resided consistently in middle latitudes and 0.65 (95% CI 0.39 to 1.18) for women who lived consistently in southern latitudes. In contrast with UC, there did not appear to be a significantly lower risk of CD according to latitude based only upon residence at birth or age 15.

We considered the possibility that a diagnosis of CD or UC might have influenced the recall of information about state of residence for those women diagnosed prior to the assessment of state of residence (1992 in NHS I and 1993 in NHSII). We therefore performed a sensitivity analysis restricting follow-up to the period after 1992 for NHS I and 1993 for NHS II. We found that, compared with participants who lived in northern latitudes, the HRs for residents in southern latitudes were 0.47 (95% CI 0.27 to 0.83) for CD and 0.65 (95% CI 0.41 to 1.03) for UC.

We explored the possibility that disease incidence varied according to latitude due to geographical variation in the prevalence of smoking (table 5). The effect of latitude of residence at age 30 on the risk of CD or UC did not appear to differ according to strata of smoking ($p_{\text{interaction}}=0.26$ and 0.99 for CD and UC, respectively). We also explored the possibility that differential patterns of diagnosis by geographical region might explain the observed association (eg, individuals may be less likely to receive a formal diagnosis of CD or UC in certain geographical locations). Thus, we examined the effect of latitude of residence

Table 3 Risk of ulcerative colitis according to latitude of residence

| | Northern latitudes | Middle latitudes | Southern latitudes |
|-------------------------------------|--------------------|---------------------|---------------------|
| Residence at birth | | | |
| Cases/person-year | 119/1 340 615 | 153/1 692 126 | 24/352 231 |
| Age-adjusted (95% CI) | 1.00 | 1.00 (0.78 to 1.27) | 0.69 (0.44 to 1.07) |
| MV-adjusted (95% CI)* | 1.00 | 1.02 (0.80 to 1.29) | 0.69 (0.44 to 1.07) |
| Residence at age 15 years | | | |
| Cases/person-year | 127/1 332 062 | 152/1 707 396 | 27/375 611 |
| Age-adjusted (95% CI) | 1.00 | 0.92 (0.72 to 1.16) | 0.67 (0.44 to 1.01) |
| MV-adjusted (95% CI)* | 1.00 | 0.93 (0.73 to 1.18) | 0.67 (0.44 to 1.01) |
| Residence at age 30 years † | | | |
| Cases/person-year | 117/1 254 794 | 148/1 636 163 | 36/537 409 |
| Age-adjusted (95% CI) | 1.00 | 0.94 (0.74 to 1.20) | 0.65 (0.45 to 0.95) |
| MV-adjusted (95% CI)* | 1.00 | 0.96 (0.75 to 1.22) | 0.64 (0.44 to 0.93) |
| Both at birth and age 15 years | | | |
| Cases/person-year | 112/1 226 893 | 141/1 572 623 | 21/288 804 |
| Age-adjusted (95% CI) | 1.00 | 0.96 (0.75 to 1.24) | 0.72 (0.45 to 1.14) |
| MV-adjusted (95% CI)* | 1.00 | 0.98 (0.76 to 1.26) | 0.72 (0.45 to 1.15) |
| Both at age 15 years and 30 years | | | |
| Cases/person-year | 107/1 080 283 | 132/1 426 019 | 20/295 840 |
| Age-adjusted (95% CI) | 1.00 | 0.91 (0.71 to 1.18) | 0.60 (0.37 to 0.97) |
| MV-adjusted (95% CI)* | 1.00 | 0.93 (0.72 to 1.20) | 0.59 (0.37 to 0.96) |
| At birth, age 15 years and 30 years | | | |
| Cases/person-year | 95/1 008 156 | 122/1 330 285 | 16/236 265 |
| Age-adjusted (95% CI) | 1.00 | 0.95 (0.73 to 1.25) | 0.64 (0.37 to 1.08) |
| MV-adjusted (95% CI)* | 1.00 | 0.97 (0.74 to 1.27) | 0.63 (0.37 to 1.08) |

*Adjusted for age (month), body mass index (<21, 21–24.9, 25–29.9, ≥30 kg/m²), ancestry (Southern European, Scandinavian, other Caucasian, non-white), oral contraceptive use (never, past, current), hormonal replacement therapy (never, past, current, premenopause) and smoking (never, past, present).

†For this analysis, cases diagnosed prior to age 30 years were excluded. MV, multivariate.

Table 4 Risk of Crohn's disease according to latitude of residence

| | Northern latitudes | Middle latitudes | Southern latitudes |
|-------------------------------------|--------------------|---------------------|---------------------|
| Residence at birth | | | |
| Cases/person-year | 102/1 340 615 | 118/1 692 126 | 26/352 231 |
| Age-adjusted (95% CI) | 1.00 | 0.91 (0.70 to 1.19) | 1.02 (0.66 to 1.57) |
| MV-adjusted (95% CI)* | 1.00 | 0.91 (0.70 to 1.19) | 1.02 (0.66 to 1.58) |
| Residence at age 15 years | | | |
| Cases/person-year | 103/1 332 062 | 118/1 707 396 | 24/375 611 |
| Age-adjusted (95% CI) | 1.00 | 0.89 (0.68 to 1.16) | 0.87 (0.55 to 1.36) |
| MV-adjusted (95% CI)* | 1.00 | 0.89 (0.68 to 1.16) | 0.86 (0.55 to 1.36) |
| Residence at age 30 years† | | | |
| Cases/person-year | 103/1 254 794 | 113/1 636 163 | 21/537 409 |
| Age-adjusted (95% CI) | 1.00 | 0.84 (0.64 to 1.09) | 0.49 (0.31 to 0.79) |
| MV-adjusted (95% CI)* | 1.00 | 0.83 (0.64 to 1.09) | 0.47 (0.27 to 0.83) |
| Both at birth and age 15 years | | | |
| Cases/person-year | 99/1 226 893 | 111/1 572 623 | 19/288 804 |
| Age-adjusted (95% CI) | 1.00 | 0.87 (0.66 to 1.14) | 0.85 (0.52 to 1.40) |
| MV-adjusted (95% CI)* | 1.00 | 0.87 (0.66 to 1.14) | 0.85 (0.51 to 1.39) |
| Both at age 15 years and 30 years | | | |
| Cases/person-year | 92/1 080 283 | 97/1 426 019 | 14/295 840 |
| Age-adjusted (95% CI) | 1.00 | 0.79 (0.59 to 1.05) | 0.59 (0.33 to 1.04) |
| MV-adjusted (95% CI)* | 1.00 | 0.79 (0.59 to 1.05) | 0.58 (0.33 to 1.02) |
| At birth, age 15 years and 30 years | | | |
| Cases/person-year | 89/1 008 156 | 92/1 330 285 | 13/236 265 |
| Age-adjusted (95% CI) | 1.00 | 0.78 (0.58 to 1.04) | 0.66 (0.37 to 1.19) |
| MV-adjusted (95% CI)* | 1.00 | 0.77 (0.58 to 1.04) | 0.65 (0.36 to 1.18) |

*Adjusted for age (month), body mass index (<21, 21–24.9, 25–29.9, ≥30 kg/m²), ancestry (Southern European, Scandinavian, other Caucasian, non-white), oral contraceptive use (never, past, current), hormonal replacement therapy (never, past, current, premenopause), smoking (never, past, present).

†For this analysis, cases diagnosed prior to age 30 years were excluded.
MV, multivariate.

according to age of diagnosis or subcategories of clinical presentation, which would be expected to vary if diagnostic criteria were not reasonably consistent across the USA. However, for UC, we did not observe significant differences in the median age at the time of diagnosis or extent of colonic involvement according to latitude of residence (table 6). Similarly, the pattern of anatomical involvement and disease

behaviour were not significantly different for cases of CD according to latitude of residence (table 6).

DISCUSSION

In two large prospective cohorts of US women, the incidence of UC and CD was significantly lower among women who resided in the southern latitudes, particularly in later life (age 30 years),

Table 5 Latitude of residence at age 30 years and risk of Crohn's disease and ulcerative colitis according to strata of smoking†

| | Northern tier | Middle tier | Southern tier | P _{Interaction} |
|-----------------------|---------------|---------------------|---------------------|--------------------------|
| Crohn's disease | | | | |
| Current/past smoker | | | | |
| Cases/person-year | 69/675 610 | 61/758 849 | 8/229 753 | 0.26 |
| Age-adjusted (95% CI) | 1.00 | 0.79 (0.56 to 1.12) | 0.36 (0.17 to 0.75) | |
| MV-adjusted (95% CI)* | 1.00 | 0.78 (0.55 to 1.10) | 0.34 (0.16 to 0.70) | |
| Never smoker | | | | |
| Cases/person-year | 34/579 184 | 52/877 314 | 13/307 657 | |
| Age-adjusted (95% CI) | 1.00 | 0.99 (0.64 to 1.52) | 0.71 (0.38 to 1.36) | |
| MV-adjusted (95% CI)* | 1.00 | 0.97 (0.63 to 1.50) | 0.72 (0.37 to 1.37) | |
| Ulcerative colitis | | | | |
| Current/past smoker | | | | |
| Cases/person-year | 71/675 610 | 77/758 849 | 17/229 753 | 0.99 |
| Age-adjusted (95% CI) | 1.00 | 0.92 (0.67 to 1.28) | 0.62 (0.36 to 1.06) | |
| MV-adjusted (95% CI)* | 1.00 | 0.90 (0.65 to 1.25) | 0.59 (0.34 to 1.00) | |
| Never smoker | | | | |
| Cases/person-year | 46/579 184 | 71/877 314 | 19/307 657 | |
| Age-adjusted (95% CI) | 1.00 | 1.02 (0.70 to 1.48) | 0.73 (0.42 to 1.24) | |
| MV-adjusted (95% CI)* | 1.00 | 1.01 (0.70 to 1.47) | 0.70 (0.40 to 1.20) | |

*Adjusted for age (month), BMI (<21, 21–24.9, 25–29.9, ≥30 kg/m²), ancestry (Southern European, Scandinavian, other Caucasian, non-white), oral contraceptive use (never, past, current), hormonal replacement therapy (never, past, current, pre-menopause), smoking (never, past, present).

†For these analyses, cases diagnosed prior to age 30 years were excluded.
MV, multivariate.

Table 6 Anatomical location and disease characteristics of women with Crohn's disease (CD) and ulcerative colitis (UC) in Nurses' Health Study (NHS) I and II

| | Northern latitude | Middle/south latitude | p Value |
|----------------------------|-------------------|-----------------------|---------|
| Ulcerative colitis (n=332) | | | |
| Age range (median), years | 28–77 (49) | 27–77 (48) | 0.63 |
| Anatomical region, % | | | 0.70 |
| Proctitis | 24 (20.7) | 37 (19.4) | |
| Left-sided | 46 (39.7) | 85 (44.5) | |
| Pan-colonic | 46 (39.6) | 69 (36.1) | |
| Crohn's disease (n=284) | | | |
| Age range (median), years | 30–80 (55) | 28–79 (49) | <0.01 |
| Anatomical region, % | | | 0.49 |
| Small bowel | 28 (27.2) | 44 (29.0) | |
| Colon | 53 (51.5) | 63 (41.4) | |
| Small bowel and colon | 22 (21.4) | 45 (29.6) | |
| Disease behaviour, % | | | |
| Stricturing (%) | 19 (18.5) | 27 (17.5) | 0.87 |
| Fistulising (%) | 12 (11.7) | 17 (11.0) | 1.00 |

Information about disease characteristics and anatomical location were obtained from the time of first report of UC or CD.

than in those residing in the northern latitudes. These results were consistent even after accounting for differences in self-reported ancestry and smoking, suggesting that other environmental or lifestyle factors correlated with geographical variation may mediate these associations. A leading explanation for this 'north–south' gradient in the risk of UC and CD may be differences in exposure to sunlight or UVB radiation, which is generally greater in southern latitudes. UV radiation is the greatest environmental determinant of plasma vitamin D and there is substantial experimental data supporting a role for vitamin D in the innate immunity and regulation of inflammatory response.^{13–14} The role of vitamin D in the pathogenesis of IBD is further supported by the observation that animal models of colitis have more severe inflammation in vitamin D receptor knock out animals or animals deficient in 1,25(OH)₂ vitamin D.^{15–18} UV radiation also induces regulatory T cells,¹⁹ and promotes production of interleukin (IL)-4 and IL-10 and inhibits production of IL-12, suppressing the inflammatory response.²⁰ Alternatively, differences in exposure to ambient air pollutants according to latitude may account for geographical variation in the risk of UC and CD.^{21–23}

Our data are supported by a number of other studies. First, an analysis based on a national health insurance database has demonstrated a clear north–south gradient for CD but not UC in France.⁵ Second, in a US veterans database, Sonnenberg and colleagues showed a higher incidence of hospital admissions associated with CD in northern states compared with southern states.⁷

Our study has several strengths distinguishing it from previous studies. First, our assessment of geographical location of residence at birth, age 15 and age 30 permitted an opportunity to examine latitude in relation to risk of CD or UC using time points with more plausible biological latency rather than relying on location at the time of diagnosis. Second, our prospective study design avoids the potential recall and selection biases of retrospective case–control studies which collect data on diet and lifestyle after diagnosis of CD or UC. Third, we confirmed all cases of CD and UC through medical record review, a significant advantage over studies that rely on self-report or discharge codes which may not accurately reflect true diagnoses. Last, our large and geographically diverse population with more than 25 years

of high follow-up rate offers the opportunity to estimate more precisely the age-adjusted incidence rate of UC or CD according to the latitude of residence.

We acknowledge several limitations. First, it is possible that our associations are related to differences in likelihood of receiving a formal diagnosis of UC or CD rather than true differences in disease incidence. However, our long-term follow-up over 25 years minimises the likelihood that true cases of UC or CD would remain undiagnosed, particularly among health professionals with a high knowledge about health with less variability in access to care. Moreover, disease behaviour or the extent of anatomical involvement did not vary according to geographical location. This suggests that there were no substantial differences in clinical presentation leading to a formal diagnosis that are likely to account for our associations. Second, although we had information on race and specific ancestry among Caucasians, it is possible that our findings are due to differences in genetic background that vary according to the latitude of residence. Nonetheless, because risk loci for CD or UC are only incompletely penetrant, it is more likely that our identification of geographical variation in the incidence of IBD still reflects at least some contribution of the environment or a complex interplay of genes and environment that mediate the risk of IBD. Third, our cohort is composed entirely of female health professionals, most of whom are Caucasian, and therefore may not represent the overall US population. However, our age-specific incidence of CD and UC were largely similar to rates from other US populations (figures 1 and 2). In addition, previous studies have shown that risk factors such as smoking and body mass index are quite consistent with those of the broader population of US women.^{24–25} Nonetheless, we acknowledge that unmeasured factors that may be uniquely associated with our cohort may limit the generalisability of our findings. Fourth, despite more than 90% follow-up over 25 years and participants' health literacy as nurses, it is possible that some cases of UC and CD were not reported. However, such misclassification of the outcome would tend to attenuate our risk estimates towards the null. Last, although a strength of our analysis is the ability to examine information about latitude of residence at birth, age 15 and age 30, we acknowledge that our analysis does not account for possible migration to different latitudes at other time points. However, in sensitivity analyses limiting the cohorts to individuals who consistently lived in the same latitude up to age 30, we observed similar effect estimates. In addition, misclassification of latitude of residence would be expected to be non-differential, attenuating our results towards the null.

In conclusion, we demonstrate that geographical residence in southern latitudes of the USA is significantly associated with a lower risk of incident CD and UC compared with residence in northern latitudes. Further studies on underlying lifestyle and environmental factors that mediate this association as well as their interaction with known genetic risk factors for CD and UC are warranted.

Acknowledgements We would like to thank the participants and staff of the Nurses' Health Study for their valuable contributions as well as the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, WY.

Funding Funded by National Institute of Health R01 CA137178, P01 CA55075, P01 CA87969, DK043351, and the Broad Medical Research Program of The Broad Foundation. ATC is a Damon Runyon Cancer Research Foundation Clinical Investigator. HK is supported by a career development award from the IBD Working Group and the Crohn's and Colitis Foundation of America. ANA is supported by a career development award from the American Gastroenterological Association.

Competing interests ATC has previously served as a consultant for Bayer Healthcare and Millennium Pharmaceuticals.

Patient consent This study is based on two large ongoing prospective studies of NHS and NHS II. These participants gave consent to be part of these studies in 1976 and 1989, respectively.

Ethics approval Ethics approval was obtained from the Institutional Review Board at the Brigham and Women's Hospital.

Contributors HK: study concept and design; acquisition of data; analysis and interpretation of data; drafting of the manuscript; statistical analysis. ESH: acquisition of data; analysis and interpretation of data; critical revision of the manuscript. ANA: acquisition of data; critical revision of the manuscript. LH: acquisition of data; analysis and interpretation of data; critical revision of the manuscript for important intellectual content. JMR: acquisition of data; analysis and interpretation of data; critical revision of the manuscript. CSF: acquisition of data; critical revision of the manuscript. ATC: study concept and design; analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Requests for access to data, statistical code, questionnaires and technical processes may be made by contacting the corresponding author at achan@partners.org.

REFERENCES

1. **Franke A**, McGovern DP, Barrett JC, *et al*. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat Genet* 2010;**42**:1118–25.
2. **Anderson CA**, Boucher G, Lees CW, *et al*. Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nat Genet* 2011;**43**:246–52.
3. **Armitage EL**, Aldhous MC, Anderson N, *et al*. Incidence of juvenile-onset Crohn's disease in Scotland: association with northern latitude and affluence. *Gastroenterology* 2004;**127**:1051–7.
4. **Green C**, Elliott L, Beaudoin C, *et al*. A population-based ecologic study of inflammatory bowel disease: searching for etiologic clues. *Am J Epidemiol* 2006;**164**:615–23; discussion 24–8.
5. **Nerich V**, Monnet E, Etienne A, *et al*. Geographical variations of inflammatory bowel disease in France: a study based on national health insurance data. *Inflamm Bowel Dis* 2006;**12**:218–26.
6. **Shivananda S**, Lennard-Jones J, Logan R, *et al*. Incidence of inflammatory bowel disease across Europe: is there a difference between north and south? Results of the European Collaborative Study on Inflammatory Bowel Disease (EC-IBD). *Gut* 1996;**39**:690–7.
7. **Sonnenberg A**, McCarty DJ, Jacobsen SJ. Geographic variation of inflammatory bowel disease within the United States. *Gastroenterology* 1991;**100**:143–9.
8. **Hernan MA**, Olek MJ, Ascherio A. Geographic variation of MS incidence in two prospective studies of US women. *Neurology* 1999;**53**:1711–18.
9. **Vieira VM**, Hart JE, Webster TF, *et al*. Association between residences in U.S. northern latitudes and rheumatoid arthritis: a spatial analysis of the Nurses' Health Study. *Environ Health Perspect* 2010;**118**:957–61.
10. **Willett W**, Stampfer MJ, Bain C, *et al*. Cigarette smoking, relative weight, and menopause. *Am J Epidemiol* 1983;**117**:651–8.
11. **Herrinton LJ**, Liu L, Lewis JD, *et al*. Incidence and prevalence of inflammatory bowel disease in a Northern California managed care organization, 1996-2002. *Am J Gastroenterol* 2008;**103**:1998–2006.
12. **Loftus CG**, Loftus EV Jr, Harmsen WS, *et al*. Update on the incidence and prevalence of Crohn's disease and ulcerative colitis in Olmsted County, Minnesota, 1940-2000. *Inflamm Bowel Dis* 2007;**13**:254–61.
13. **Lim WC**, Hanauer SB, Li YC. Mechanisms of disease: vitamin D and inflammatory bowel disease. *Nat Clin Pract Gastroenterol Hepatol* 2005;**2**:308–15.
14. **Cantorna MT**, Zhu Y, Froicu M, *et al*. Vitamin D status, 1,25-dihydroxyvitamin D₃, and the immune system. *Am J Clin Nutr* 2004;**80**:1717S–20S.
15. **Froicu M**, Cantorna MT. Vitamin D and the vitamin D receptor are critical for control of the innate immune response to colonic injury. *BMC Immunol* 2007;**8**:5.
16. **Froicu M**, Zhu Y, Cantorna MT. Vitamin D receptor is required to control gastrointestinal immunity in IL-10 knockout mice. *Immunology* 2006;**117**:310–18.
17. **Froicu M**, Weaver V, Wynn TA, *et al*. A crucial role for the vitamin D receptor in experimental inflammatory bowel diseases. *Mol Endocrinol* 2003;**17**:2386–92.
18. **Cantorna MT**, Munsick C, Bemiss C, *et al*. 1,25-Dihydroxycholecalciferol prevents and ameliorates symptoms of experimental murine inflammatory bowel disease. *J Nutr* 2000;**130**:2648–52.
19. **Maeda A**, Beissert S, Schwarz T, *et al*. Phenotypic and functional characterization of ultraviolet radiation-induced regulatory T cells. *J Immunol* 2008;**180**:3065–71.
20. **Ullrich SE**. Sunlight and skin cancer: lessons from the immune system. *Mol Carcinog* 2007;**46**:629–33.
21. **Kaplan GG**, Hubbard J, Korzenik J, *et al*. The inflammatory bowel diseases and ambient air pollution: a novel association. *Am J Gastroenterol* 2010;**105**:2412–19.
22. **Ananthakrishnan AN**, McGinley EL, Binion DG, *et al*. Ambient air pollution correlates with hospitalizations for inflammatory bowel disease: an ecologic analysis. *Inflamm Bowel Dis* 2011;**17**:1138–45.
23. **Beamish LA**, Osornio-Vargas AR, Wine E. Air pollution: an environmental factor contributing to intestinal disease. *J Crohns Colitis* 2011;**5**:279–86.
24. **Chiuve SE**, Fung TT, Rexrode KM, *et al*. Adherence to a low-risk, healthy lifestyle and risk of sudden cardiac death among women. *JAMA* 2011;**306**:62–9.
25. **Sarna L**, Bialous SA, Jun HJ, *et al*. Smoking trends in the Nurses' Health Study (1976-2003). *Nurs Res* 2008;**57**:374–82.