Ambient air quality as risk factor for microscopic colitis – A geographic information system (GIS) study

Bas P.M. Verhaegh, Esmee M. Bijnens, Tim R.A. van den Heuvel, Danny Goudkade, Maurice P. Zeegers, Tim S. Nawrot, Ad A.M. Masclee, Daisy M.A.E. Jonkers, Marieke J. Pierik

Keywords:
Microscopic colitis
Epidemiology
Risk factors
Environmental factors
Air pollution

Keywords:

Contents lists available at ScienceDirect

Environmental Research

journal homepage: www.elsevier.com/locate/envres

Introduction

Microscopic colitis (MC) is a chronic bowel disorder with frequent watery diarrhoea as primary symptom. Although the aetiology is not clear, MC is considered a multifactorial disease, in which immunological, genetic, microbial, life style and environmental factors play a role (Pisani et al., 2016). Amongst the latter, smoking and drug exposure have repeatedly been associated with MC (Yen et al., 2012; Verhaegh et al., 2016; Masclee et al., 2015; Jaruvongvanich et al., 2019; Burke et al., 2018; Wickbom et al., 2017). Especially (co)exposure to non-steroidal anti-inflammatory drugs (NSAIDs) and proton pump inhibitors (PPIs) seems to increase the risk of MC (Verhaegh et al., 2016; Masclee et al., 2015). However, in these studies only 20–40% of MC patients were exposed to these drugs in the year before diagnosis, suggesting that drug-exposure does not explain for all MC cases. Other possible risk factors for MC (e.g. hormones, alcohol, socioeconomic status) have been addressed in a limited number of studies, but the results were either negative or conflicting (Yen et al., 2012; Roth et al., 2013, 2014; Sonnenberg et al., 2017; Pabla and Ness, 2018). Moreover, there are indications for an altered microbiome in MC patients, compared to healthy controls (Fischer et al., 2015; Rindom Krogsbaard et al., 2019). In other words, further research for possible
risk factors is warranted in order to confirm current risk factors and hypotheses on underlying pathophysiological mechanisms, or to find new ones.

Several publications report that smoking is associated with MC (Yen et al., 2012; Jaruvongvanich et al., 2019; Burke et al., 2018; Fernandez-Baneres et al., 2013). Tobacco smoke compounds such as particulate matter (PM), nitric oxides and benzene are also present in polluted air. Therefore, ambient air quality (influenced by e.g. industrial, agricultural and traffic related emissions in the residential area) may have a contributory effect to the pathogenesis of MC as well. Furthermore, MC peak incidence rates are observed in the older population (> 60 years of age), which might also be supportive for involvement of environmental factors in the pathogenesis to some extent. Considering that people are continuously exposed to polluting sources in the direct environment, such as traffic, industry, urbanized areas or agricultural poisons, ambient air quality might be contributory. Moreover, previous studies reported an association between air pollution and various gastrointestinal conditions, such as gastrointestinal cancer, appendicitis, bowel infections and IBD (Garcia-Perez et al., 2010; Kaplan et al., 2009, 2010; Orazzo et al., 2009; Opstelten et al., 2016). However, the association with MC has never been reported.

Chemicals present in ambient air can reach the gastrointestinal tract by direct ingestion or via ingestion of pulmonary mucus, which pairing epithelial permeability (Salim et al., 2014). In addition, some bacteria are able to metabolize ingested particles into toxic metabolites affecting mucosal barrier function. In general, environmental factors (in)directly induce epigenetic changes during life. In a genetic susceptible host, the sum of epigenetic changes and environmental influences then may trigger inflammatory activity, not resolving due to the epigenetic changes (Rogler and Vavricka, 2015).

In conclusion, there is circumstantial evidence which suggests that ambient air quality may play a role in MC pathophysiology. Therefore, the aim of this study was to assess the effect of ambient air quality on the risk of MC in a population-based cohort of MC patients. We hypothesized that exposure to ambient air pollution increases the risk of MC.

2. Materials and methods

2.1. Study population

A case-control study was performed, including all MC cases in the region of South Limburg, the Netherlands. Cases were identified in PALGA, the Dutch nationwide registry of histo- and cytopathology. All cases aged 18 years or older, with a PALGA registered diagnosis of MC, collagenous colitis (CC), lymphocytic colitis (LC), incomplete CC (CCi) or incomplete LC (LCi) between January 2000 and December 2012, were selected. Hereafter, pathology reports and medical charts were reviewed to exclude cases which did not meet the histological or clinical criteria for MC. The haematoxylin-eosin stained sections from diagnostic biopsies were retrieved and revised according to the ESP/EMCG diagnostic criteria (Langner et al., 2015) by an experienced pathologist (D.G.), in order to verify the diagnosis. The pathologist was blinded for any clinical symptoms and the previously established diagnosis.

Non-MC controls were retrieved from the general South Limburg population. Study information letters were sent to randomly selected households within the study area, based on postal code. Per household, informed consent was asked to a maximum of four subjects above 18 years of age. Of this control population (n = 1,611), two non-MC controls were matched to each MC case, based on gender and year of birth (± 2 years). If a control was matched, any other member from the same household was excluded for further matching. Frequency matching was applied to ensure that the control population resembled the distribution of the area’s background population.

The index date of the cases was defined as the date of histological diagnosis. Non-MC controls were assigned the same index date as their matched case.

2.2. Residential data

For each subject, historical and current residential data were retrieved from the national civil registration system (in Dutch: Basisregistratie Persoonsgegevens, https://www.rvig.nl/brp), kept by the Dutch Ministry of Internal Affairs. For each subject, the residential address at index date was recorded. Only subjects with a stable residential address for at least three years before the index date, were considered suitable for inclusion. Residential addresses were geocoded, based on the data provided by the ‘Basisregistratie Adressen en Gebouwen’, a database maintained by the national Cadastre survey (https://www.kadaster.nl/bag).

2.3. Markers for ambient air quality

Both direct and indirect markers for ambient air quality were included in the study, i.e. main air pollution components, land use data, population density, urbanity, proximity to major roads, and total road length in the residential area. Based on the geocoded residential addresses, individual exposure to the various variables was assessed, applying geographic information system (GIS) functions (ArcGIS 9.3. Esri, Redlands, CA, USA). Fig. 1 (panel C–F) gives an impression on the spatial distribution of some of the included markers for ambient air quality.

Data on regional concentrations of the most common air pollutants according to the World Health Organisation, i.e. particulate matter ≤10 μm (PM10), nitric dioxide (NO2), ozone (O3), sulphur dioxide (SO2) and benzene (C6H6), were obtained from the National Institute for Public Health (RIVM) of the Dutch Ministry of Health, Welfare and Sports (http://www.lml.rivm.nl). Air pollutant concentrations (μg/m3) were measured at 60 fixed monitoring stations in the Netherlands. Measured concentrations, added with local and foreign data, were applied in a dispersion model to calculate local concentrations on a 5 × 5km grid and interpolated to a 1 × 1km grid (RIVM, 2015). A margin of uncertainty of 15% around the interpolated data was calculated by the RIVM. Data were available for each year between 2000 and 2012. In the study analyses, SO2 and benzene were not included, as rational concentrations were too low to expect any health risks (RIVM, 2015).

Land use data were retrieved from CORINE Land Cover, a cartographic database (scale 1:100,000) coordinated by the European Environment Agency (EEA) (European). Each 100 × 100 m was assigned to its major type of land use, being residential (CORINE class 1.1), industrial (class 1.2.1), urban green (class 1.4.1), agricultural (class 2) or natural areas (class 3). Data were available for the years 2000, 2006, and 2012. Within 100 m, 500 m and 2500 m radius buffers from each residential address, the percentages of the various types of land use were determined. Data were not restricted to national borders. Traffic intensity was reflected by a) the total road length within 100 m, 200 m, 500 m radius buffers from the residential address and b)
Fig. 1. Panel A and B show the distribution of the included cases and controls in South Limburg. Panels C–F visualize the spatial distribution of particulate matter (PM) (C), land use (D), major roads (E) and population density (F) in South Limburg, for the year 2006. These panels give an impression of the source data and do not represent the complete data, as multiannual data were used for the analyses.
### Table 1

<table>
<thead>
<tr>
<th></th>
<th>Cases (n = 345)</th>
<th>Controls (n = 583)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female, n (%)</td>
<td>271 (78.6%)</td>
<td>450 (77.2%)</td>
<td>0.66</td>
</tr>
<tr>
<td>Age at index date, mean years ± SD</td>
<td>63.3 ± 13.0</td>
<td>61.2 ± 12.4</td>
<td>0.01</td>
</tr>
<tr>
<td>Duration of stable address before index date, mean years ± SD</td>
<td>21.3 ± 12.8</td>
<td>21.3 ± 12.1</td>
<td>0.96</td>
</tr>
<tr>
<td>Type of MC – PA revision</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC, n (%)</td>
<td>108 (31.3%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LC, n (%)</td>
<td>176 (51.0%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CGi, n (%)</td>
<td>6 (1.7%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LCi, n (%)</td>
<td>28 (8.1%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unrevised, n (%)</td>
<td>15 (4.3%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smoking status at index date</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current, n (%)</td>
<td>151 (43.8%)</td>
<td>89 (15.3%)</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>Former, n (%)</td>
<td>81 (23.5%)</td>
<td>236 (40.5%)</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>Never, n (%)</td>
<td>113 (32.8%)</td>
<td>258 (44.3%)</td>
<td>&lt; 0.01</td>
</tr>
</tbody>
</table>

**CC**: Collagenous colitis; **LC**: lymphocytic colitis; **CCI**: incomplete collagenous colitis; **LCi**: incomplete lymphocytic colitis.

*Diagnosis was based on the conclusion provided in the pathology report.

### 2.4. Statistical analysis

Continuous variables were expressed as means with standard deviation (SD). With regard to the patient characteristics, statistical significant differences between cases and controls were calculated using an independent Student's T-test for continuous variables or a Chi²-test for categorical variables. For all other variables, univariate logistic regression analysis was applied to calculate odds ratios (OR) with 95% Confidence Intervals (95%-CI). All variables with a p-value < 0.10 in the univariate analyses were included in the multivariable logistic regression analysis. Both univariate and multivariate models were corrected for the confounders age, gender and smoking status at index date. In the case population, smoking status at index date was obtained by scrutinizing patient's medical files. Controls were asked for their smoking status (including year of cessation, if applicable) via a brief questionnaire added to the informed consent form. Missing data on smoking status were imputed by applying a fully conditional specification method in SPSS. This method imputes missing data with random values from a multinomial distribution based on the category proportions. Presence of multicollinearity between selected variables was assessed by requesting collinearity diagnostics in the statistical program. Statistical significance was determined as p < 0.05.

For all analyses, CCI and LCi cases were included in the CC and LC group, respectively. Furthermore, in case biopsy specimens could not be revised, the diagnosed subtype as recorded in the pathology report was acquired, provided that the diagnosis was set with a high or moderate likelihood (Verhaegh et al., 2015). A sensitivity analysis was performed, excluding CCI and LCi cases and cases without a biopsy proven diagnosis.

All analyses were conducted with IBM SPSS Statistics version 20.0 (SPSS Inc., Chicago, IL, USA).

### 2.5. Ethical considerations

This study was performed according to the revised version of the declaration of Helsinki and approved by the medical ethical committee of Maastricht University Medical Centre* (NL4127.068.13 and NL31636.068.10). Written informed consent was obtained from all participants.

### 3. Results

#### 3.1. Study population

In total, 430 MC cases were selected from the three participating centres, of which 77 cases were excluded for the following reasons: they were living outside the designated study area at index date (n = 5), they lived < 3 years on the residential address since at index date (n = 60), or they were included in two centres at the same time (n = 2). In addition, 18 cases were excluded because the diagnosis could not be confirmed upon biopsy revision.

The remaining 345 cases (271 females, 78.6%) consisted of 108 (31.1%) CC, 176 (51.0%) LC, 6 CCI (1.7%) and 28 (8.1%) LCi cases. In 27/345 cases (7.8%, 15 CC and 12 LC), no biopsy material was available for revision (Table 1). As all of them had a high likelihood of a positive MC diagnosis based on the pathology report, they were included for further analyses.

All 345 cases were matched to 583 non-MC controls (450 females, 77.2%). The spatial distribution of the included subjects is visualized in Fig. 1 (panel A and B). A 1:2 matching ratio was achieved in 69.0% of cases. As presented in Table 1, the mean age at index date was 63.3 ± 13.0 years for the cases and 61.1 ± 12.4 years for the controls (p = 0.01). On average, the duration of an unchanged residential address before the index date was 21.3 ± 12.8 and 21.3 ± 12.1 years (p = 0.96) for cases and controls, respectively.

In total, the smoking status at index date was missing in 13.0% of the cases and 4.3% of the controls. After imputation, the proportion of current smokers was 43.8% in the case population and 15.3% in the control population (p < 0.01) (Table 1).

#### 3.2. Environmental factors

##### 3.2.1. Univariate analyses

Data on the various measured (proxy) markers for ambient air quality were presented in Table 2. Only the percentage of urban green within a 500 m radius buffer around the residential address (OR 0.20; 95%-CI 0.03–1.27) and residential proximity to the nearest highway (OR 0.93; 95%-CI 0.87–1.00) were statistically different between cases and controls on a p < 0.10 level. No statistical differences were observed between cases and controls regarding the percentage of industrial, residential, agricultural, natural or urban green area, total road length around the residential address, population density, address density, or major air pollution compounds (Table 2).

##### 3.2.2. Multivariable analysis

Next to the percentage of urban green within a 500 m radius buffer and residential proximity to the nearest highway; gender, age and smoking status at index date were included in the multivariable model. Based on this model, a higher age at diagnosis (OR 1.02, 95%-CI 1.01–1.04) and current smoking at index date (OR 4.30; 95%-CI 3.01–6.14) were significantly associated with a higher risk of MC (Table 3).

##### 3.2.3. Sensitivity analysis

All analyses were repeated after exclusion of CCI/LCi cases and...
4. Discussion

This is the first study to explore the role of ambient air quality on the risk of MC. Although various direct and indirect markers for ambient air quality were studied, none of them were significantly associated with an increased risk of MC on a multivariable level. Though, the previously reported risk factors age and current smoking were confirmed.

Ambient air is a heterogeneous composition of gases, particles and volatile organic compounds (VOCs) (e.g. benzene). Especially PM10 and NO2, emitted by both natural and anthropogenic sources (e.g. industry, traffic, agriculture), have a negative impact on general health by increasing the risk of mortality, hospital admissions and pulmonary, cardiovascular or gastrointestinal disorders (Beamish et al., 2011; Heroux et al., 2015). However, the results of the present study showed that the primary components and proxy markers for air pollution (e.g. PM10, high industrial area, residential proximity to roads) were not associated with MC on the univariate and multivariable level, which makes a major role for ambient air quality in MC pathophysiology unlikely. However, a possible contributory effect should not be discarded completely. Although the level of exposure to ambient air pollution was not different between MC cases and controls, the effect of this exposure might still be different between the two populations, for instance due to (genetic) susceptibility. As an example, in inflammatory bowel disease the -gene is only associated with Crohn’s disease in a smoking population (van der Heide et al., 2010), which supports the hypothesis that specific genetic variants are required for specific environmental factors to contribute to disease development.

In contrast to most (proxy) markers for ambient air quality, current smoking was found to be an incontestable risk factor for MC, which is in line with various recent studies (Yen et al., 2012; Jaruvongvanich et al., 2009).
2019; Burke et al., 2018; Fernandez-Banares et al., 2013; Munch et al., 2016). Considering the fact that there is overlap in the composition of cigarette smoke and polluted air, the results of this study might be an argument that other compounds of cigarette smoke, generally not present in polluted air (e.g. nicotine) do account for the strong associations. However, it should be noted that the level of exposure to toxic gases and particles via cigarette smoke is considerably higher and more frequent in comparison to ambient air pollution.

The strength of this study resides in the ability to link the subject’s residential address at index date to outcome variables from the same area and time period, which increased the validity of the results. Although no eminent role for ambient air quality was observed in this study, similar approaches linking ambient air quality data to health outcomes were able to detect significant associations by applying GIS (Rutten et al., 2017; Nawrot et al., 2011; Bijjens et al., 2016; Pinichka et al., 2017; Albi and Mirzahosseini, 2013), also in the same area (Munch et al., 2016). In addition, the case population in the current study was population-based, well-defined, relatively large, and with a revised histological diagnosis in 92% of the cases. Furthermore, controls were derived from the same area, were adequately distributed over the study area and were solidly matched to the MC cases. Although the control population was two years younger on average, we do not think this had major impact on the study outcome. Last, analyses have been corrected for a major MC risk factor, i.e. smoking status at index date. Despite 4–13% of data was missing, residual confounding was reduced by imputation of these missing data.

Legal limits for average annual concentrations of PM<sub>10</sub> and NO<sub>2</sub> in the Netherlands are 40 μg/m<sup>3</sup> and WHO guideline values are 20 μg/m<sup>3</sup> for PM<sub>10</sub> and 40 μg/m<sup>3</sup> for NO<sub>2</sub> (WHO, 2016). The concentrations in our study area were generally below these values. This means that the results of this study are conclusive for low exposed populations only. Therefore, the conclusions need to be confirmed in different populations, preferably with higher exposure to the assessed risk factors.

Some limitations should be noted. First, the size of the study area may have been too small to have sufficient variation in the distribution of air pollution determinants. Second, the sample size may have been too small to detect small variations in the exposure variables. Third, the residential address is often not the location where people spend most of their day (work, traveling, time spend in traffic). Therefore, exposure to environmental factors on the primary residential address may not be representative for an individual’s level of exposure to air pollution. Fourth, the source data applied in this study were not available for each index year. Although most recent data relative to the index date were applied, and three-year average concentrations for air pollution compounds were calculated, small annual variations in the exposure variables might have influenced the outcomes. Finally, it should be noted that data on drug exposure at index date (e.g. NSAIDs or PPIs) were not available.

In conclusion, this was the first study to explore an association between ambient air quality and the risk of MC. Based on the results of this study, ambient air quality is not a major contributing factor to the risk of MC.

Conflicts of interest

None to be declared.

Funding

None to be declared.

Authorship statement

Guarantor of the article: Bas Verhaegh takes responsibility for the integrity of the work as a whole, from inception to published article.

Specific author contributions: BV, EB, TH, DJ designed the work and performed the data analyses, data interpretation and data acquisition, TN contributed to the data analyses and data interpretation. All authors participated in the draft and/or revision of the manuscript. All read and approved the final draft submitted.

Acknowledgements

The establishment of this study was supported by the help of the following people, who facilitated either access to the local patient files or provision of local pathology slides: L. Oostenbrug, M. Romberg-Camps, P. Sastrowijoto, R. Clarjīs (Zuyderland MC, Heerlen/Sittard).

References


Pabla, B., Niss, R.M., 2018. Do sex hormones cause, or are they only associated with, microscopic colitis? Gastroenterology 155 (6), 1679–1681.


RIVM, 2015. Grootschalige Concentratie en Depositiekaarten Nederland. RIVM.

Rogler, G., Vavricka, S., 2015. Exposome in IBD: recent insights in environmental factors that influence the onset and course of IBD. Inflamm. Bowel Dis. 21 (2), 400–408.


Roth, B., Gustafsson, R.J., Jeppsson, B., Manjer, J., Ohlsson, B., 2014. Smoking- and alcohol habits in relation to the clinical picture of women with microscopic colitis compared to controls. BMC Women’s Health 14, 16.


