Research article

JHEP|Reports

Geo-epidemiology and environmental co-variate mapping of primary biliary cholangitis and primary sclerosing cholangitis

Jessica Katharine Dyson,[1,2,*] Alasdair Blain,[3] Mark David Foster Shirley,[4] Mark Hudson,[1,2] Steven Rushton,[4] David Emrys Jeffreys Jones[1,2]

[1]*Translational and Clinical Research Institute, Newcastle University, Newcastle-upon-Tyne, UK;* [2]*Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle-upon-Tyne, UK;* [3]*Wellcome Centre for Mitochondrial Research, Newcastle University, Newcastle-upon-Tyne, UK;* [4]*School of Natural and Environmental Sciences, Newcastle University, Newcastle-upon-Tyne, UK*

JHEP Reports **2021.** https://doi.org/10.1016/j.jhepr.2020.100202

Background & Aims: Autoimmune liver disease (AILD) is thought to result from a complex interplay between genetics and the environment. Studies to date have focussed on primary biliary cholangitis (PBC) and demonstrated higher disease prevalence in more urban, polluted, and socially deprived areas. This study utilises a large cohort of patients with PBC and primary sclerosing cholangitis (PSC) to investigate potential environmental contributors to disease and to explore whether the geo-epidemiology of PBC and PSC are disease-specific or pertain to cholestatic AILD in general.
Methods: All adult patients with PBC and PSC in a tightly defined geographical area within the UK were identified. Point- and area-based analyses and structural equation modelling (SEM) were used to investigate for disease clustering and examine for relationships between prevalence, distribution of environmental contaminants, and socio-economic status.
Results: We identified 2,150 patients with PBC and 472 with PSC. Significant spatial clustering was seen for each disease. A high prevalence of PBC was found in urban, post-industrial areas with a strong coal-mining heritage and increased environmental cadmium levels, whereas a high PSC prevalence was found in rural areas and inversely associated with social deprivation.
Conclusions: This study demonstrates spatial clustering of PBC and PSC and adds to our understanding of potential environmental co-variates for both diseases. Disease clustering, within the same geographical area but over different scales, is confirmed for each disease with distinct risk profiles identified and associations with separate putative environmental factors and socio-economic status. This suggests that different triggers and alternative pathways determine phenotypic expression of autoimmunity in the affected population. Co-variate analysis points towards the existence of specific disease triggers.

Introduction

Autoimmune disease is thought to arise as a result of a complex interaction between genetic susceptibility factors and environmental triggers.[1] There has been substantial progress in recent years in identifying genetic contributors (typically the cumulative impact of numerous individually low impact, typically immune-related, loci). There has been considerably less progress, however, in identifying environmental triggers and why disease develops in only a small proportion of genetically predisposed individuals remains a key unanswered question.

The current literature regarding the geo-epidemiology of, and potential environmental factors in, autoimmune liver disease (AILD) has focussed mainly on primary biliary cholangitis (PBC) with identification of significant spatial and spatio-temporal variations in disease risk. Studies from the north-east of England have identified a higher prevalence of PBC in urban areas,[2–4] whereas a study in New York City found clusters of PBC patients in zip codes that contained, or were adjacent to, superfund toxic waste sites (SFSs).[5] These studies point to the importance of potential environmental risk factors in disease aetiology but the same level of investigation of primary sclerosing cholangitis (PSC), the other cholestatic AILD, and autoimmune hepatitis (AIH) have not previously been undertaken. The only modelling study in PSC was as part of the work from New York that did not show any statistically significant difference between the prevalence of transplanted PSC patients according to zip code or for type of SFS.[5] In AIH, a single study found clusters of patients listed for liver transplantation near sites with high levels of chlorinated hydrocarbons including trichloroethylene.[6]

PBC and PSC are both rare, but important, causes of immune-mediated, cholestatic, chronic liver disease and account for up to 18% of elective adult liver transplants in the UK.[7] There is strong evidence for a genetic component in both diseases, with genome-wide association studies and other studies identifying a number of susceptibility loci and risk genes.[8–11]

Keywords: Primary biliary cholangitis; Primary sclerosing cholangitis; Autoimmune hepatitis; Geo-epidemiology; Socio-economic status; Cadmium; Urban; Rural.
Received 8 June 2020; received in revised form 14 September 2020; accepted 9 October 2020Available online 4 November 2020
* Corresponding author. Address: Liver Unit, Freeman Hospital, High Heaton, Newcastle-upon-Tyne, England, UK, NE7 7DN. Tel.: +44 191 244 8606.
E-mail address: Jessica.dyson1@nhs.net (J.K. Dyson).

In this study, we have used more advanced modelling techniques than previous studies to build on the existing geo-epidemiological research in PBC. In addition to investigating disease clustering we have used point-based analyses, area-based analyses, and structural equation modelling (SEM) to attempt to understand environmental triggering by introducing spatial covariates (risk factors) into the models. This study assesses both PBC and PSC as comparator cholestatic AILDs to evaluate whether putative environmental factors are common to cholestasis or specific to each disease. We have then used AIH as a comparator non-cholestatic, but immune-mediated, liver disease to explore the hypothesis that potential environmental triggers are disease-specific rather than features shared by autoimmunity *per se*. We believe that the findings make a substantial contribution to our understanding of disease pathogenesis and demonstrate distinct risk profiles for the development of PBC in comparison with PSC. The techniques used here have given unique insights into these diseases, and this study is an exemplar for a broadly applicable approach.

Patients and methods

Populations

A comprehensive case-finding approach was used to identify all incident and prevalent patients with PBC, PSC, and AIH in a tightly-defined geographical study area (the Academic Health Science Network for the North East and North Cumbria [AHSN NENC]) who had been diagnosed before the end of 2016. This area was used as the denominator as it reflects natural clinical referral patterns.

Diagnostic criteria

Patients with 'definite' or 'probable' PBC, based on the conventional diagnostic criteria, were included.[12,13] The standard diagnostic criteria for PSC were used; cholestatic liver biochemistry with typical cholangiographic features in the absence of other identifiable causes or other extrahepatic disease (except inflammatory bowel disease) with patients with clinical, biochemical, and histological features compatible with PSC, but with a normal cholangiogram, being classified as small duct PSC.[14] The diagnosis of AIH was based on the simplified International AIH Group diagnostic criteria from 2008.[15] Patients with anti-mitochondrial antibody positivity and normal liver blood tests (n = 694), those with an *a priori* clinical diagnosis of an overlap syndrome (PBC/AIH: 130 patients; PSC/AIH: 36 patients), those with an unknown postcode or living in a postcode outside the study area (n = 299) and patients who were not alive during the time period studied (n = 793) were excluded from analyses.

Multiple case-finding methodologies were used: World Health Organization International Statistical Classification of Diseases and Related Health Problems 10 codes, liver histopathology reports (Table S1), autoantibody profiles, radiology reports, and outpatient clinic letters. The lead researcher, a consultant hepatologist with a specialist interest in AILD, reviewed all case notes to confirm or refute the clinical diagnosis. Further details of the case-finding methodologies used are provided in the Supplementary data: Case-finding and search strategies.

Ethical approval was obtained on 12 March 2015 (REC reference 15/SW/0048, IRAS project ID 166616, ISRCTN48732084). In compliance with the Declaration of Helsinki, NHS research and development approval was obtained for each study site before recruitment. The study was conducted in accordance with International Conference on Harmonisation Good Clinical Practice guidance. Valid, written informed consent was obtained in accordance with the study protocol from participants for whom additional clinical data were collected.

Modelling methods

Point-based analyses

Further details of the modelling methods used are provided in the Supplementary data. Spatial point-based analyses (K function) were used to investigate for the presence and patterns of clustering of disease. K function analyses were used to assess if cases occurred together in space and time more than expected by chance. This method counts the number of cases within concentric rings of each case and then compares the observed count with that expected as a random process.[16] Place of residence (postcode) was used as the spatial reference point and year of diagnosis was used for temporal modelling.

'Postcode head counts' from the NOMIS dataset were used to create a pool of control sites equivalent to the total population of the region (approximately 3 million people) to adjust for population size. Modelling included adjustment for the size of each areal unit and population density.

Area-based analyses

Point-based analyses do not identify the size or location of clusters or account for the role of putative risk factors. Area-based analyses were used to examine for relationships between disease prevalence and the distribution of potential environmental contaminants and socio-economic status. The null hypothesis was that there was no difference between areas in expected disease prevalence given the population size and distribution of potential environmental contaminants. Bayesian area-based analyses (using conditional autoregressive models) were used to estimate the relative risk of disease in individual postcode districts. The expected numbers of cases were calculated by distributing the total observed cases for the study area between the postcode districts according to the adult population estimate for each postcode district, that is NE3, CA7 (part of the postcode classification system in the UK). Models were then fitted with a variety of explanatory variables as single spatial covariates. Coal mines, landfill sites, limestone quarries, sandstone quarries, and lead mines were all corrected for area, that is counts per unit area. The deviance information criterion (DIC) was used to compare model fit with 97.5% confidence intervals. A difference in DIC>2 indicates a significant difference in model performance.[17,18] The best performing model was that with the lowest DIC containing only significant variables.

Structural equation modelling

Taken together, point- and area-based analyses of patterns of disease allow us to generate hypotheses about the putative causes of disease. However, they do not take into account additional factors that may interact to contribute to disease development. Our hypothesis was that putative environmental and socio-economic risk factors are themselves inter-related. Therefore, structural equation modelling (SEM) was used to develop a conceptual model (Fig. 1) with pathways linking factors or covariates that are hypothesised to be driven by or drive disease itself. Potential sources of environmental exposure were divided into 3 types:

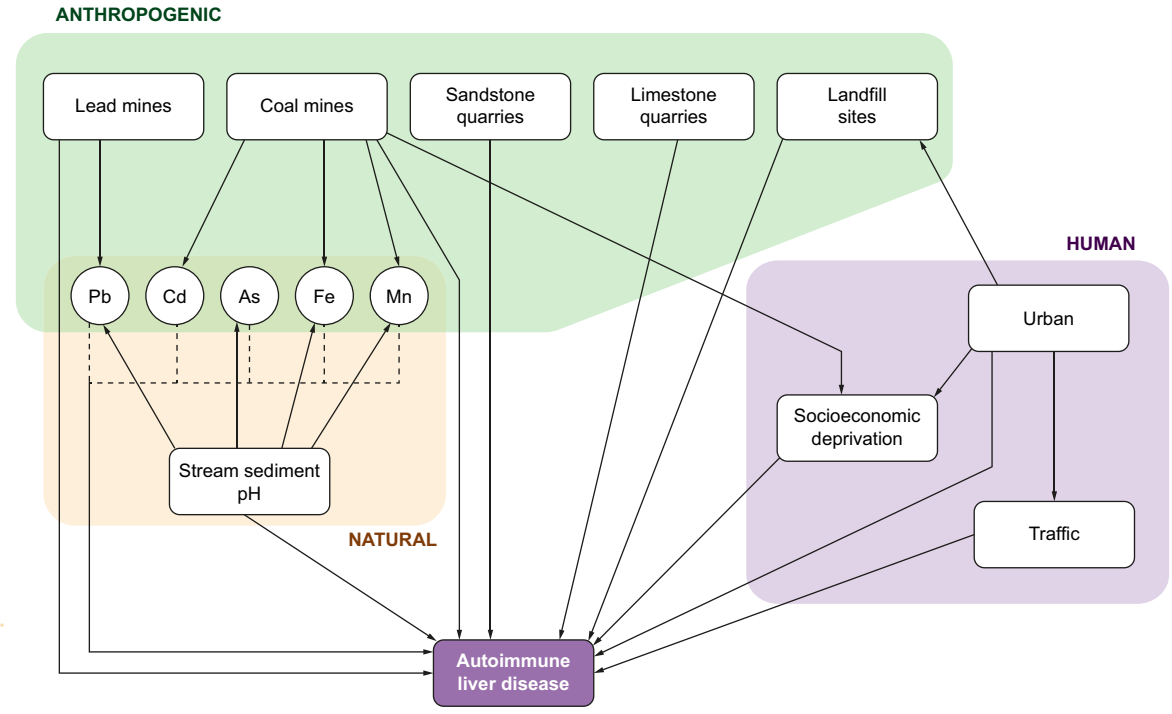

Fig. 1. Conceptual structural equation model for autoimmune liver disease.

1. Arising from the natural environment – landscape features and putative toxins,
2. Anthropogenic sources of pollutants,
3. Human influences on the environment resulting in population exposure.

The SEM used actual counts of diseases and was adjusted for population size. Model fit was estimated using the root mean square error of association (RMSEA) and comparative fit index (CFI). Non-significant variables were removed from the model and it was re-run until the best model was identified in which all included covariates were significant at the 95% level.

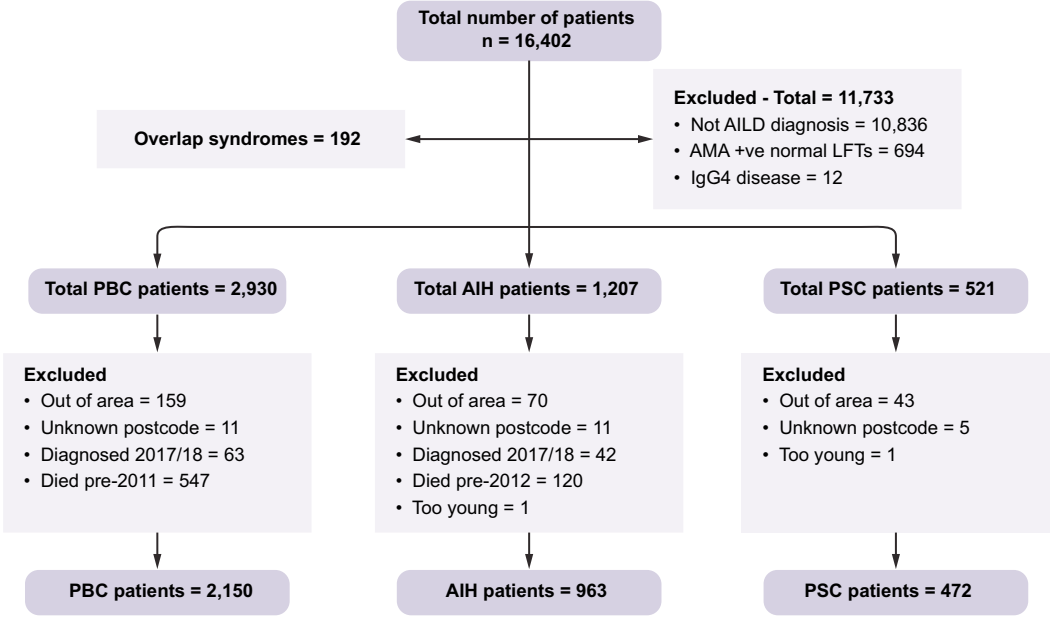

Fig. 2. Consolidated Standards of Reporting Trials diagram. AIH, autoimmune hepatitis; AILD, autoimmune liver disease; AMA, anti-mitochondrial antibody; PBC, primary biliary cholangitis; PSC, primary sclerosing cholangitis.

Spatial covariates

Hypotheses were generated regarding potential environmental triggers related to autoimmune liver disease. 'Urban-ness' was assessed using the Rural-Urban Classification, giving a measure of the proportion of people living in an urban environment with an average value for each postcode district. Traffic count datasets for 2000–2015 were obtained from the Department for Transport for the north-east and north-west of England. Data regarding landfill sites, coal mines, lead mines, sandstone quarries, and limestone quarries were obtained from the Environment Agency Geostore and the British Geological Society BRITPITS Database (licence number 2016/076BP ED). The Geochemical Baseline Survey of the Environment project provided stream sediment pH and heavy metal data. The Townsend Material Deprivation Score and The Index of Multiple Deprivation (IMD) were used as measures of social deprivation.

Non-parametric data are presented as median and range. Continuous variables are described as median, minimum, and maximum and assessed using a 2-tailed Mann–Whitney *U* test with $p < 0.05$ being considered statistically significant. Data were analysed using SPSS version 22 (SPSS Inc., Chicago, IL, USA), GraphPad Prism 7 (GraphPad Software, San Diego, CA, USA) and in the R statistical package, version 3.6 (R Foundation for Statistical Computing, Vienna, Austria).

Results

There were 2,150 PBC patients (1906; 88.7% female, point prevalence 41.7/100,000 population) and 472 PSC patients (151; 32% female, point prevalence 8.6/100,000 population) identified with residential postcodes within the study area, giving a total study population of 2,622 patients. For the comparator AIH group, a further 963 patients (772; 80.2% female, point prevalence 21.2/100,000 population) were identified. Fig. S1 shows the distribution of the cases for the individual diseases across the study area. Fig. 2 shows a Consolidated Standards of Reporting Trials diagram outlining the patient identification process, reasons for excluding patients and the final number of patients available for inclusion in the study.

Primary biliary cholangitis

There was clear spatial clustering of PBC with the observed K-function being outside the limits of the random samples between approximately 1–2 km and then at all distances above 7.5 km (Fig. 3A). This suggests clustering at both a local and a broader level. There was no significant temporal clustering (Fig. S2A). After adjusting for population size, areas of relative high and low PBC prevalence were observed (Fig. 3B) with statistically significant variation in prevalence (Fig. 3C). The DIC for the 'null model' was 799.755. When spatial covariates were added to the model, coal mines gave the largest change in DIC (9.885) with a highly significant increase in effect, suggesting a strong association between PBC and coal mining (Table 1). Urban-ness also significantly improved the null model but with a change in DIC score of <2. The kriging maps in Fig. S3 show that coal mines map onto urban areas and those with higher levels of social deprivation. The areas with significantly increased PBC prevalence (Table S2), such as Blyth and Cramlington in Northumberland, are former coal-mining areas with more recent housing development, frequently on reclaimed mine sites.

The best SEM model for PBC is shown in Fig. 4A. The final model was a good fit with an RMSEA of 0.000 and CFI of 1.000. Coal mines

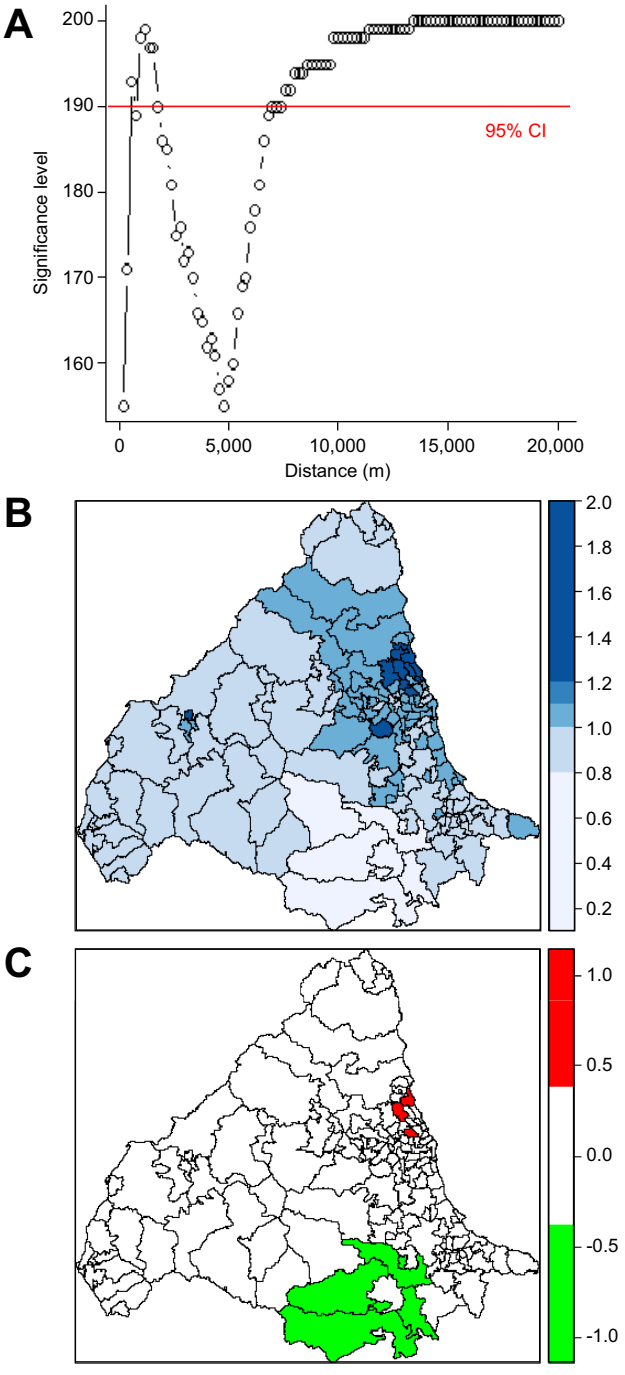

Fig. 3. K-function analysis of spatial clustering and conditional autoregressive models of relative risk for PBC across study area. (A) K-function analysis of spatial clustering for 2,150 PBC patients. Significant clustering (using 95% CI) occurring when the data lies above the red line. (B) Relative risk for PBC in each postcode district. (C) Significant relative risk for PBC in each postcode district with dichotomised 'high' (red) and 'low' (green) relative risk map at 97.5% significance. PBC, primary biliary cholangitis.

Table 1. Log file for DIC scores for null model and single covariates in PBC.

	2.50%	Median	97.50%	DIC	Change in DIC
Null				799.755	
Urban	0.006796	0.1048	0.2014	799.209	0.546
Traffic	−0.00934	−0.00333	0.002562	801.347	−1.592
Landfill sites	−0.656	0.7062	2.039	801.621	−1.866
Coal mines	1.06000	1.992	3.029	789.87	9.885
Limestone quarries	−4.1000	0.3970	4.757	801.758	−2.003
Sandstone quarries	−0.527	0.08870	0.7020	801.223	−1.468
Lead mines	−15.5000	7.4600	28.96	801.02	−1.265
Manganese	−0.1349	0.05671	0.2543	801.648	−1.893
Lead	−0.00028	0.0000749	0.000414	801.084	−1.329
Arsenic	−0.00607	−0.00156	0.002764	801.058	−1.303
Iron	−0.00713	−0.00125	0.003499	801.18	−1.425
Cadmium	−0.04128	−0.00491	0.02935	802.019	−2.264
Stream sediment pH	−0.06738	0.0141	0.0918	802.899	−3.144
Townsend score	−0.0237	0.004068	0.03112	801.563	−1.808

Red line (Coal mines): Best model – statistically significant improvement on null model with largest change in DIC score. Yellow line (Urban): Statistically significant improvement on null model but with change in DIC score <2.
DIC, deviance information criterion; PBC, primary biliary cholangitis.

were the major contributor (39%) to the SEM and environmental cadmium levels appeared to play an interactive role with coal mine distribution. Cadmium levels also made a direct contribution to the model (22%). The kriging map in Fig. 4B shows that coal mines map onto many of the cadmium-rich areas.

Primary sclerosing cholangitis

Significant spatial clustering was seen for PSC with peaks at approximately 1 km, 2 km, and between 7.5 and 12.5 km (Fig. 5A) with no significant temporal clustering (Fig. 2Fig. S2B). Areas of high and low prevalence of PSC were observed (Fig. 5B), with statistically significant variation in disease risk (Fig. 5C).

The DIC for the null model for PSC was 577.200. When spatial covariates were included (Table 2), urban-ness and Townsend score significantly improved the 'null model'. Both were negatively associated with the change in DIC score, that is fewer urban (more rural) and fewer deprived areas had a higher prevalence of PSC. The area with increased PSC prevalence is in the Lake District in Cumbria, an almost exclusively rural, sheep-farming area (Table S3).

The best SEM model for PSC only had an RMSEA of 0.141 and CFI of 0.872 with p >0.08 for all variables, so the model could not be refined any further. However, it identified a number of covariates (stream sediment pH, arsenic, lead mines, and lead) that were associated with variation in disease risk that warrant further investigation.

Autoimmune hepatitis

To contextualise the results for PBC and PSC we compared the findings in these 2 diseases to the non-cholestatic, autoimmune liver disease AIH. Statistically significant spatial clustering was seen for AIH with a peak of clustering at approximately 1 km and further clustering between 7.5 and 14 km (Fig. 6A) but with no temporal clustering (Fig. S2C). Although there was variation in relative risk of AIH across the study area, this was not statistically significant using 97.5% confidence intervals (Fig. 6B and C). The DIC for the null model for AIH was 681.621 and none of the single spatial covariates improved the model (Table S4).

For AIH, the best SEM model had an RMSEA of 0.000 and CFI of 1.000 with no p values >0.05. Coal mines contributed 6% to the variation in disease risk whereas cadmium was an independent risk factor (22%). The relationship between stream sediment pH and AIH was negative, that is there was more disease in more acidic areas.

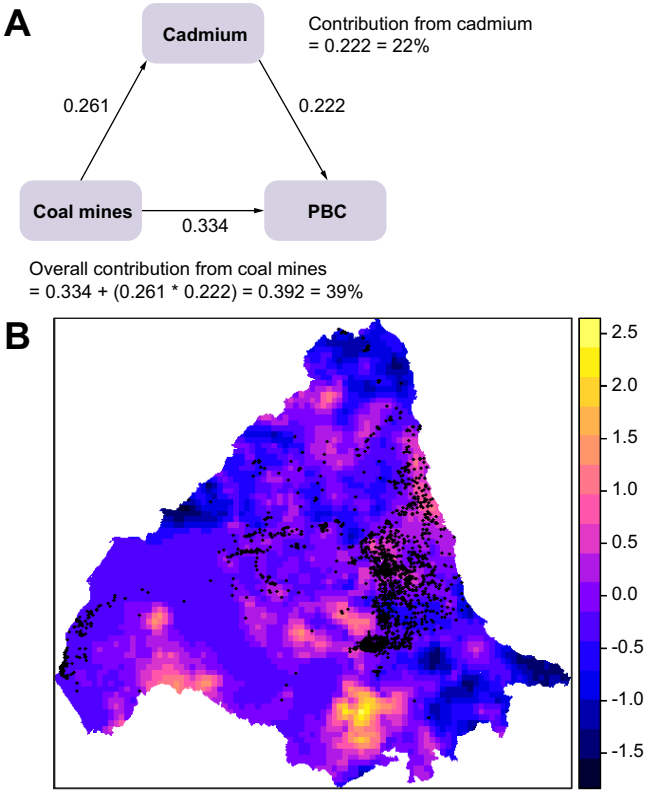

Fig. 4. Structural equation model for PBC and distribution of coal mines and environmental cadmium levels across study area. (A) Output from best SEM model for PBC. (B) Locations of coal mines (represented by black dots) overlaid onto kriging map of stream sediment cadmium concentrations. The lighter the colour on the scale, the higher the concentration of cadmium (yellow = highest, blue = lowest). PBC, primary biliary cholangitis; SEM, structural equation modelling.

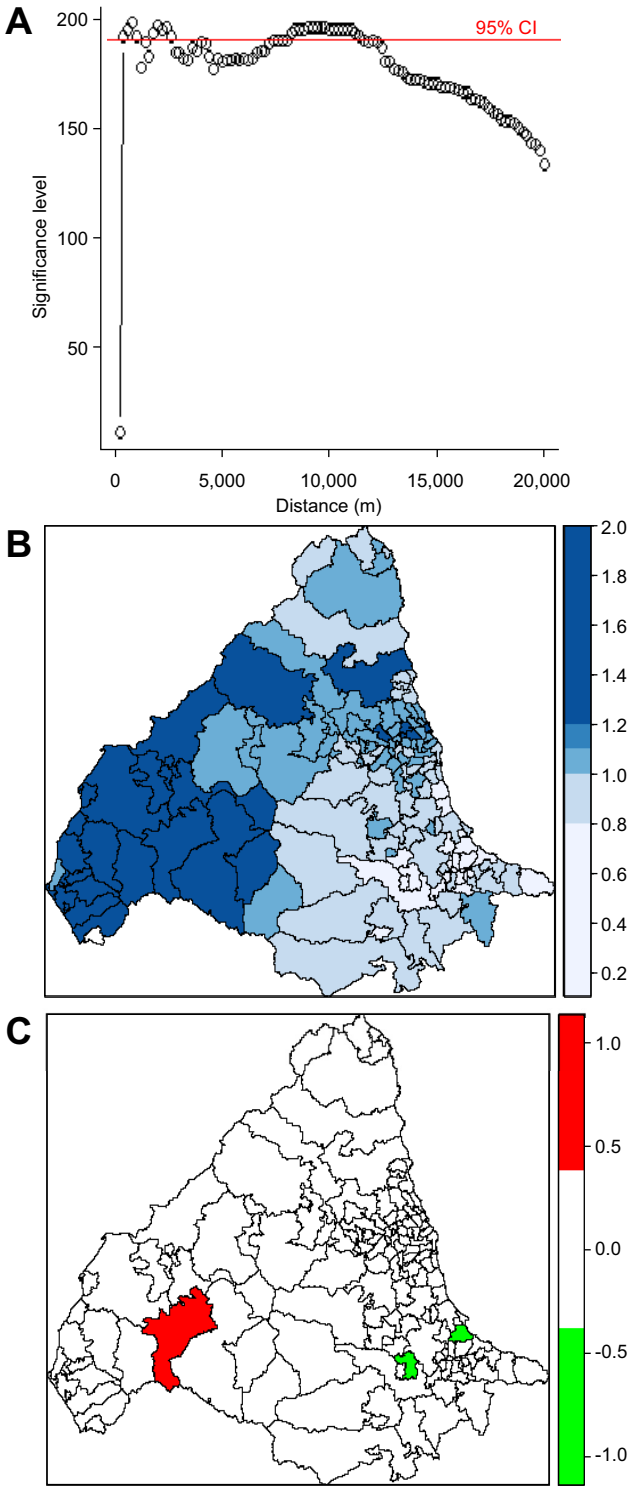

Fig. 5. K-function analysis of spatial clustering and conditional autoregressive models of relative risk for PSC across study area. (A) K-function analysis of spatial clustering for 472 PSC patients. (B) Relative risk for PSC each postcode district. (C) Significant relative risk for PSC in each postcode district with dichotomised 'high' (red) and 'low' (green) relative risk map. PSC, primary sclerosing cholangitis.

Discussion

This is one of the most comprehensive studies of the geo-epidemiology of PBC and PSC performed to date. It demonstrates the potential of comprehensive epidemiological approaches for increasing our understanding of disease aetiology through identification of potential environmental and socioeconomic risk factors. It builds on previous work carried out in the north-east of England exploring the geo-epidemiology of PBC,[2–4,19–22] but here, the identical modelling approaches have also been used to study PSC in the same geographical area. We have demonstrated spatial, area-based clustering of both PBC and PSC but with notable differences in their geographical distribution. There was a high prevalence of PBC in the urban, post-industrial east of the study region that has a strong coal-mining heritage, as opposed to PSC, which was more common in the rural, sheep-farming west and inversely associated with social deprivation. An identical modelling approach was used for AIH (as a comparator non-cholestatic, autoimmune liver disease). Although significant spatial clustering was seen, the variation in relative risk of AIH across the study area was not significant and no spatial covariates were identified. The modelling techniques used here were more advanced than in the previous north-east England PBC studies, used a much larger cohort (2,150 *vs.* 770 with minimal cohort overlap given the 25-year difference between the studies), included a broader geographical area and investigated other AILD. Previous studies aimed at identifying disease triggers in both PBC and PSC have largely used case–control epidemiological methods.[3,4,21,23–26] The previous geo-epidemiological study from New York that did include both PBC and PSC patients found an association with toxic landfill sites for PBC but not for PSC.[5] There are potential synergies between the different experimental approaches and their findings. Tobacco smoke, landfill sites, and industrial emissions are major sources of volatile organic compounds, other aromatic hydrocarbons, polychlorinated biphenyls, and heavy metals, all of which are known to be related to immune dysregulation.[27,28]

It is known that there is an increased risk of PBC and PSC in the family members of patients compared with the general population.[29–31] Given the well-described genetic associations of both diseases, the potential for a genetic basis for this observation is clear. Parents, offspring, and siblings typically have a shared environment for many years, raising the possibility that familial risk may also have an environmental component. Point-based analyses showed peaks at short distances (1–2 km) and then a second, larger peak at a greater distance (above 7.5 km). The first peak could potentially relate to a family with more than one member affected by disease. Alternatively, small postcodes/areas are likely to have less heterogeneity in many parameters (*e.g.* genetics, potential exposure risks) that could result in close case clustering. The presence of 2 peaks may point to genetics being one factor in disease aetiology (accounting for peaks over small distances) with the second peak relating to broader environmental associations.

There was no temporal clustering in any of the diseases studied. This is in contrast to previous studies from the north-east of England that found statistically significant space–time clustering of PBC patients.[2,3] The authors suggested that a transient environmental cause (such as an acute infection) originating from a fixed geographical source could contribute to disease development. It is perhaps a surprising discovery in a chronic disease that has different modes of presentation, and where there is an undefined, and often prolonged, time period

Table 2. Log file for DIC scores for null model and single covariates in PSC.

	2.50%	Median	97.50%	DIC	Change in DIC
Null				577.200	
Urban	−0.3424	−0.1704	−0.0012	574.892	2.308
Traffic	−0.01624	−0.00413	0.0077	578.302	−1.102
Landfill sites	−4.4400	−1.444	1.325	577.405	−0.205
Coal mines	−2.0000	0.1076	2.249	578.722	−1.522
Lead mines	−19.300	18.00	48.98	577.677	−0.477
Sandstone quarries	−1.4200	−0.094300	1.2200	579.573	−2.373
Limestone quarries	−9.9200	−0.70400	7.751	576.885	0.315
Cadmium	−0.1054	−0.02616	0.04727	577.931	−0.731
Arsenic	−0.00563	0.00299	0.01126	578.816	−1.616
Lead	−0.000972	−0.000213	0.000484	579.155	−1.955
Manganese	−0.2235	0.1577	0.5254	577.322	−0.122
Iron	−0.00992	0.000427	0.009606	577.944	−0.744
Stream sediment pH	−0.2868	−0.00223	0.2175	579.162	−1.962
Townsend score	−0.1190	−0.06367	−0.0122	574.416	2.784

Orange line (Urban): Statistically significant improvement on null model with change in DIC score >2. Red line (Townsend score): Best model – statistically significant improvement on null model with largest change in DIC score.
DIC, deviance information criterion; PSC, primary sclerosing cholangitis.

between developing disease and the clinical diagnosis being made.

One of the key findings in this study is that SEM showed that up to 39% of the variation in PBC risk was associated with proximity to coal mines, and also pointed to relationships with environmental cadmium both as an independent covariate and also as an additive risk factor to residence in a former coal-mining location. The SEM for PSC could not be refined to the same degree but showed different covariates to that seen in PBC (stream sediment pH, arsenic, lead mines, and lead). Possible hypotheses about why coal mining may be related to PBC include environmental disturbance leading to release of potential xenobiotics, products from processing escaping into the environment, a surrogate association, socio-economic aspects (*e.g.* smoking rates), and communities with similar genetics being challenged by another pathology that stimulates the immune response resulting in disease development.

Cadmium exposure can cause liver injury[32] and is a well-recognised environmental pollutant with cadmium mobilisation as a result of coal-mining activity and water pollution from abandoned metal mines recognised in environmental policy.[33–35] The traditional conceptual model for PBC pathogenesis is one of a primary autoimmune disease with the initial injury to biliary epithelial cells (BECs) being a consequence of breakdown of immune self-tolerance. More recently, an alternative hypothesis has emerged in which the primary injury to the BECs is cytopathic, with altered expression of self-antigen (pyruvate dehydrogenase) as a consequence of injury.[36] Cadmium is cytopathic, inducing apoptosis, cell senescence, and epithelial to mesenchymal transition[37]; all cardinal features of the BEC response to injury in PBC.[32,38–40]

The observation of PSC being more common in rural, affluent areas raises questions regarding the inter-play between environmental entities in the landscape and socio-economic status. Overall, rural areas tend to be less deprived than urban ones. Government statistics show that “12% of people living in urban areas are in areas that are within the most deprived 10% of the IMD, compared with just 1% of people living in rural areas”.[41] However, owing to the IMD being a measure of relative deprivation, not every individual in a deprived area will be deprived, and *vice versa*. The inverse relationship between PSC and social deprivation is in stark contrast to most diseases where prevalence is higher in areas of greater social deprivation with poorer levels of overall health.[42–44]

The link between PSC and rurality could reflect land usage, for example agricultural activity, use of pesticides, and/or fertilisers. Organophosphates are used to improve agricultural yield but related to the development of autoimmune diseases in humans.[45] Studies suggesting that genetic polymorphisms may be relevant in determining individual response to pesticide exposure point to the interaction between genetics and the environment.[46] Work in rats has shown that concomitant exposure to organophosphates increases hepatic reactive oxygen species and triggers an acute liver injury with rises in transaminases.[47]

Limitations and future work

One of the main limitations in both the previous and current studies is the use of date of diagnosis as a surrogate for date of disease onset. This means that conclusions regarding a temporal element to disease onset should be interpreted with caution.

The current work examines associations between PBC and PSC and the environment at a population level. Individual-level data (*e.g.* smoking status, employment) should now be used to examine these relationships at that scale and investigate for the presence of additional aetiological factors contributing to disease clustering. The choice of covariates was based on hypotheses regarding potential environmental triggers but it must be acknowledged that the covariates included were dependent on the data being available at the geographical unit studied.

One of the challenges in epidemiological research relates to patient mobility. Individuals may move within the course of their lifetime and space–time clustering may be affected by population shifts during the study period. However, it is well-reported that the north of England has low migrations rates, particularly in women aged 30 years or over (almost all patients with PBC).[22] In the current study, for the 1,450 patients for whom both the postcode at diagnosis and postcode at the time of study entry were available, 73% were the same at postcode unit level (*i.e.* in the same house or within the same cluster of houses), 85% at postcode district level and 98% at postcode area level. This confirms that although changing residential address was a limitation within the study, there was relatively little movement within the study region.

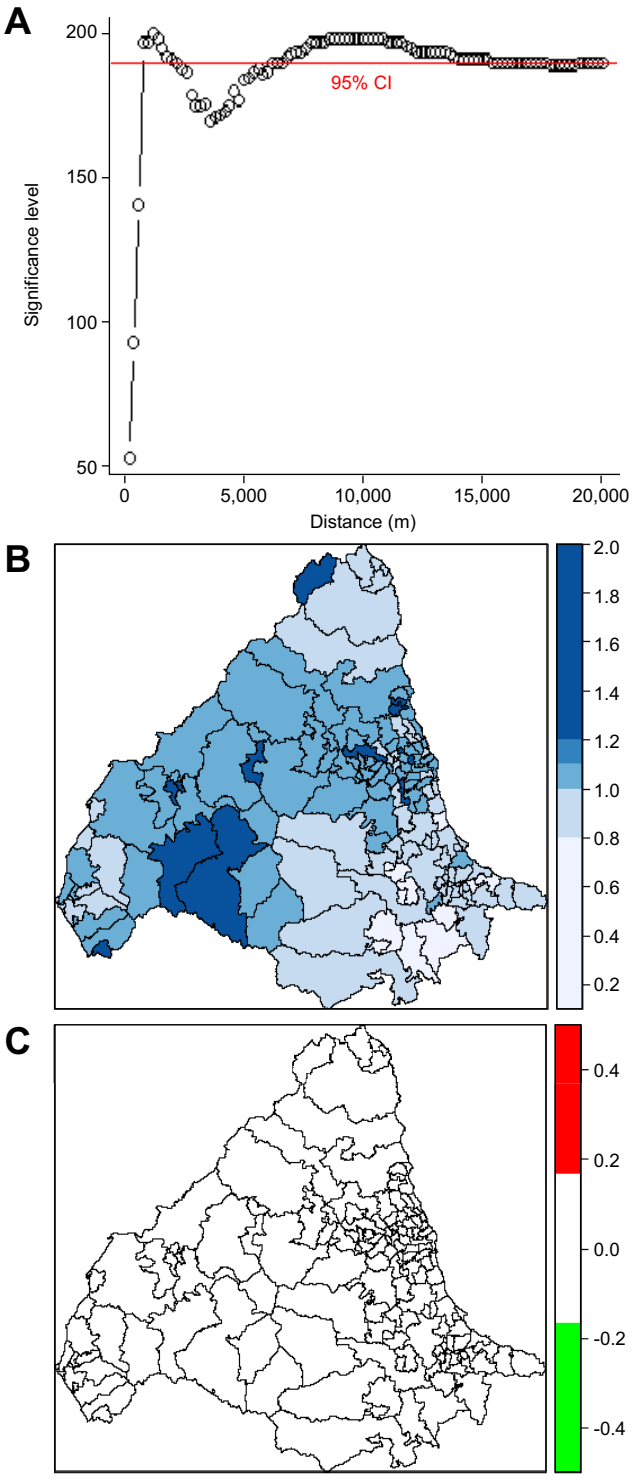

Fig. 6. K-function analysis of spatial clustering and conditional autoregressive models of relative risk for AIH across study area. (A) K-function analysis of spatial clustering for 963 AIH patients. (B) Relative risk for AIH each postcode district. (C) Significant relative risk for AIH in each postcode district with dichotomised 'high' (red) and 'low' (green) relative risk map. AIH, autoimmune hepatitis.

Conclusions

Association and correlation do not imply causality. However, this study has demonstrated novel findings of disease clustering and associations with putative environmental risk factors that are different between PBC and PSC. The distinct risk profiles associated with each disease have not previously been reported and add significantly to the current literature. This work suggests that there may be a common predisposition (such as genetics) in the affected population with different triggers and alternative pathways determining the phenotypic expression of autoimmunity. Improved understanding of disease pathogenesis may enable reduced exposure to or modification of the effect of potential triggers on susceptible individuals.

Abbreviations

AHSN NENC, Academic Health Science Network for the North East and North Cumbria; AIH, autoimmune hepatitis; AILD, autoimmune liver disease; BECs, biliary epithelial cells; CFI, comparative fit index; DIC, deviance information criterion; IMD, Index of Multiple Deprivation; PBC, primary biliary cholangitis; PSC, primary sclerosing cholangitis; RMSEA, root mean square error of association; SEM, structural equation modelling; SFS, superfund toxic waste site.

Financial support

JKD is funded by the National Institute for Health Research (NIHR) Rare Diseases Translational Research Collaboration and the Newcastle NIHR Biomedical Research Centre (BH149219/PD0252). The NIHR had no role in study design, data collection/analysis/interpretation, report writing or the decision to submit the paper for publication.

Conflict of interest

The authors declare no conflicts of interest that pertain to this study.

Please refer to the accompanying ICMJE disclosure forms for further details.

Authors' contributions

Study design: JKD, MH, DEJJ, SR. Literature search and data collection: JKD. Data analysis and interpretation: JKD, MS, AB, SR. Wrote the initial version of the manuscript: JKD, DEJJ. Reviewed and approved the final version of the manuscript: all authors.

Data availability

Owing to the detailed geographical information collected about patients, participants were assured that raw data would remain confidential and would not be shared.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.jhepr.2020.100202.

References

[1] Zenewicz LA, Abraham C, Flavell RA, Cho JH. Unraveling the genetics of autoimmunity. Cell 2010;140:791–797.

[2] McNally RJ, Ducker S, James OF. Are transient environmental agents involved in the cause of primary biliary cirrhosis? Evidence from space–time clustering. Hepatology 2009;50:1169–1174.

[3] McNally RJ, James PW, Ducker S, James OF. Seasonal variation in the patient diagnosis of primary biliary cirrhosis: further evidence for an environmental component to etiology. Hepatology 2011;54:2099–2103.
[4] McNally RJ, James PW, Ducker S, Norman PD, James OF. No rise in incidence but geographical heterogeneity in the occurrence of primary biliary cirrhosis in North East England. Am J Epidemiol 2014;179:492–498.
[5] Ala A, Stanca CM, Bu-Ghanim M, Ahmado I, Branch AD, Schiano TD, et al. Increased prevalence of primary biliary cirrhosis near superfund toxic waste sites. Hepatology 2006;43:525–531.
[6] Stanca CM, Babar J, Singal V, Ozdenerol E, Odin JA. Pathogenic role of environmental toxins in immune-mediated liver diseases. J Immunotoxicol 2008;5:59–68.
[7] NHS Blood and Transplant. Organ Donation and Transplantation Activity Report 2017/18. Available at, https://nhsbtdbe.blob.core.windows.net/umbraco-assets-corp/12300/transplant-activity-report-2017-2018.pdf. [Accessed September 2019].
[8] Hirschfield GM, Liu X, Han Y, Gorlov IP, Lu Y, Chen W, et al. Variants at IRF5-TNPO3, 17q12-21 and MMEL1 are associated with primary biliary cirrhosis. Nat Genet 2010;42:655–657.
[9] Ji SG, Juran BD, Mucha S, Folseraas T, Jostins L, Melum E, et al. Genome-wide association study of primary sclerosing cholangitis identifies new risk loci and quantifies the genetic relationship with inflammatory bowel disease. Nat Genet 2016;49:269–273.
[10] Liu JZ, Hov JR, Folseraas T, Ellinghaus E, Rushbrook SM, Doncheva NT, et al. Dense genotyping of immune-related disease regions identifies nine new risk loci for primary sclerosing cholangitis. Nat Genet 2013;45:670–675.
[11] Mells GF, Floyd JA, Morley KI, Cordell HJ, Franklin CS, Shin SY, et al. Genome-wide association study identifies 12 new susceptibility loci for primary biliary cirrhosis. Nat Genet 2011;43:329–332.
[12] Hirschfield GM, Dyson JK, Alexander GJM, Chapman MH, Collier J, Hubscher S, et al. The British Society of Gastroenterology/UK-PBC primary biliary cholangitis treatment and management guidelines. Gut 2018;67:1568–1594.
[13] Lindor KD, Gershwin ME, Poupon R, Kaplan M, Bergasa NV, Heathcote EJ, et al. Primary biliary cirrhosis. Hepatology 2009;50:291–308.
[14] Chapman MH, Thorburn D, Hirschfield GM, Webster GGJ, Rushbrook SM, Alexander G, et al. British Society of Gastroenterology and UK-PSC guidelines for the diagnosis and management of primary sclerosing cholangitis. Gut 2019;68:1356–1378.
[15] Hennes EM, Zeniya M, Czaja AJ, Pares A, Dalekos GN, Krawitt EL, et al. Simplified criteria for the diagnosis of autoimmune hepatitis. Hepatology 2008;48:169–176.
[16] Bailey TC, Gattrell A. Interactive Spatial Data Analysis. London: Longman Scientific and Technical; 1995.
[17] Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. Bayesian measures of model complexity and fit (with discussion). J R Stat Soc 2002;64:583–639.
[18] Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. The deviance information criterion: 12 years on (with discussion). J R Stat Soc 2014;76:485–493.
[19] Metcalf JV, Mitchinson HC, Palmer JM, Jones DEJ, Bassendine MF, James OFW. Natural history of early primary biliary cirrhosis. Lancet 1996;348:1399–1402.
[20] Prince M, Chetwynd A, Newman W, Metcalf JV, James OF. Survival and symptom progression in a geographically based cohort of patients with primary biliary cirrhosis: follow-up for up to 28 years. Gastroenterology 2002;123:1044–1051.
[21] Prince MI, Ducker SJ, James OFW. Case-control studies of risk factors for primary biliary cirrhosis in two United Kingdom populations. Gut 2010;59:508–512.
[22] Prince MI, Chetwynd A, Diggle P, Jarner M, Metcalf JV, James OFW. The geographical distribution of primary biliary cirrhosis in a well-defined cohort. Hepatology 2001;34:1083–1088.
[23] Andersen IM, Tengesdal G, Lie BA, Boberg KM, Karlsen TH, Hov JR. Effects of coffee consumption, smoking, and hormones on risk for primary sclerosing cholangitis. Clin Gastroenterol Hepatol 2014;12:1019–1028.
[24] Corpechot C, Chretien Y, Chazouilleres O, Poupon R. Demographic, lifestyle, medical and familial factors associated with primary biliary cirrhosis. J Hepatol 2010;53:162–169.
[25] Eaton JE, Juran BD, Atkinson EJ, Schlicht EM, Xie X, de Andrade M, et al. A comprehensive assessment of environmental exposures among 1000 North American patients with primary sclerosing cholangitis, with and without inflammatory bowel disease. Aliment Pharmacol Ther 2015;41:980–990.
[26] Gershwin ME, Selmi C, Worman HJ, Gold EB, Watnik M, Utts J, et al. Risk factors and comorbidities in primary biliary cirrhosis: a controlled interview-based study of 1032 patients. Hepatology 2005;42:1194–1202.
[27] Crinnion WJ. Polychlorinated biphenyls: persistent pollutants with immunological, neurological, and endocrinological consequences. Altern Med Rev 2011;16:5–13.
[28] Milnerowicz H, Sciskalska M, Dul M. Pro-inflammatory effects of metals in persons and animals exposed to tobacco smoke. J Trace Elem Med Biol 2015;29:1–10.
[29] Bergquist A, Lindberg G, Saarinen S, Broome U. Increased prevalence of primary sclerosing cholangitis among first-degree relatives. J Hepatol 2005;42:252–256.
[30] Bergquist A, Montgomery SM, Bahmanyar S, Olsson R, Danielsson A, Lindgren S, et al. Increased risk of primary sclerosing cholangitis and ulcerative colitis in first-degree relatives of patients with primary sclerosing cholangitis. Clin Gastroenterol Hepatol 2008;6:939–943.
[31] Jones DE, Watt FE, Metcalf JV, Bassendine MF, James OF. Familial primary biliary cirrhosis reassessed: a geographically-based population study. J Hepatol 1999;30:402–407.
[32] Salama SA, Arab HH, Hassan MH, Al Robaian MM, Maghrabi IA. Cadmium-induced hepatocellular injury: modulatory effects of gamma-glutamyl cysteine on the biomarkers of inflammation, DNA damage, and apoptotic cell death. J Trace Elem Med Biol 2019;52:74–82.
[33] Galunin E, Ferreti J, Zapelini I, Vieira I, Ricardo Teixeira Tarley C, Abrao T, et al. Cadmium mobility in sediments and soils from a coal mining area on Tibagi River watershed: environmental risk assessment. J Hazard Mater 2014;265:280–287.
[34] Mishra VK, Upadhyaya AR, Pandey SK, Tripathi BD. Heavy metal pollution induced due to coal mining effluent on surrounding aquatic ecosystem and its management through naturally occurring aquatic macrophytes. Bioresour Technol 2008;99:930–936.
[35] De Chowdhuri R. Pollution alert – cadmium level 169 times higher than is safe found in Chhattisgarh's Raigarh district. Available at, 2017. https://scroll.in/latest/846331/cadmium-level-169-times-higher-than-is-safe-found-in-chhattisgarhs-raigarh-district-says-study. [Accessed September 2019].
[36] van Niekerk J, Kersten R, Beuers U. Role of bile acids and the biliary HCO_3^- umbrella in the pathogenesis of primary biliary cholangitis. Clin Liver Dis 2018;22:457–479.
[37] Ospondpant D, Phuagkhaopong S, Suknuntha K, Sangpairoj K, Kasemsuk T, Srimaroeng C, et al. Cadmium induces apoptotic program imbalance and cell cycle inhibitor expression in cultured human astrocytes. Environ Toxicol Pharmacol 2019;65:53–59.
[38] Jarup L, Akesson A. Current status of cadmium as an environmental health problem. Toxicol Appl Pharmacol 2009;238:201–208.
[39] Pares A, Rimola A, Bruguera M, Mas E, Rodes J. Renal tubular acidosis in primary biliary cirrhosis. Gastroenterology 1981;80:681–686.
[40] Thomas LD, Hodgson S, Nieuwenhuijsen M, Jarup L. Early kidney damage in a population exposed to cadmium and other heavy metals. Environ Health Perspect 2009;117:181–184.
[41] Ministry of Housing, Communities and Local Government. Index of multiple deprivation. 2019. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/854569/Deprivation_2019.pdf. [Accessed 3 November 2020].
[42] Keane MG, Horsfall LJ, Rait G, Pereira SP. Sociodemographic trends in the incidence of pancreatic and biliary tract cancer in UK primary care. PLoS One 2014;9:e108498.
[43] Pollock AM, Vickers N. Breast, lung and colorectal cancer incidence and survival in South Thames Region, 1987–1992: the effect of social deprivation. J Public Health Med 1997;19:288–294.
[44] Smith GD, Hart C, Watt G, Hole D, Hawthorne V. Individual social class, area-based deprivation, cardiovascular disease risk factors, and mortality: the Renfrew and Paisley Study. J Epidemiol Community Health 1998;52:399–405.
[45] Corsini E, Liesivuori J, Vergieva T, Van Loveren H, Colosio C. Effects of pesticide exposure on the human immune system. Hum Exp Toxicol 2008;27:671–680.
[46] Tabrez S, Priyadarshini M, Priyamvada S, Khan MS, Na A, Zaidi SK. Gene-environment interactions in heavy metal and pesticide carcinogenesis. Mutat Res Genet Toxicol Environ Mutagen 2014;760:1–9.
[47] Dwivedi N, Flora SJ. Concomitant exposure to arsenic and organophosphates on tissue oxidative stress in rats. Food Chem Toxicol 2011;49:1152–1159.