Muirhead CR, James OFW, Ducker SJ, McNally RJQ.

[Does primary biliary cirrhosis cluster in time?](#)

Spatial and Spatio-temporal Epidemiology 2015, 14-15, 1-8.

Copyright:

This is the author's version of a work accepted for publication. ©2015. This manuscript is made available under the [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International](#) license

DOI link to article:

<http://dx.doi.org/10.1016/j.sste.2015.06.001>

Date deposited:

09/06/2015

Embargo release date:

16 June 2016

This work is licensed under a
[Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International](#) licence

Does Primary Biliary Cirrhosis cluster in time?

Colin R. Muirhead PhD,^{1*} Oliver F.W. James FRCP,² Samantha J. Ducker², Richard J.Q.

McNally PhD¹

¹ Institute of Health & Society, Newcastle University, Newcastle upon Tyne, UK

² Institute of Cellular Medicine, Newcastle University, Newcastle upon Tyne, UK

* Corresponding author: Dr Colin R Muirhead, Institute of Health & Society, Newcastle University, Baddiley-Clark Building, Richardson Road, Newcastle upon Tyne, NE2 4AX, UK; Tel: +44-(0)191-208-3699; Fax: +44(0)191-208-6043. Email - Colin.Muirhead@Newcastle.ac.uk

Running title: Does Primary Biliary Cirrhosis cluster in time?

Keywords: Aetiology; Infection; Primary Biliary Cirrhosis; Temporal Clustering

Contributors: This work was conceived and designed by RM and CM, who both also oversaw the conduct of the study. OJ and SD were responsible for data collection. CM conducted the statistical analyses. All authors contributed to the interpretation of the analyses. The manuscript was drafted by CM and RM with input from OJ and SD. All of the authors approved the final version.

Highlights

- This is the first study to look for irregular temporal patterns in diagnoses of PBC.
- There was no evidence for general irregular (non-seasonal) temporal clustering.
- There was long-term temporal variation in PBC diagnoses.
- Agents occurring in geographically widespread mini-epidemics are not implicated.
- Future work should look for agents occurring in localised sporadic mini-epidemics.

ABSTRACT

The aetiology of primary biliary cirrhosis (PBC) is not well established. Previously we found evidence of space-time clustering and seasonal variation in the date of diagnosis, suggesting a possible role for a transient or seasonally varying environmental factor. We examined whether a temporally varying environmental agent may be involved by analysing population-based PBC data from northeast England over 1987-2003. Using an adaptation of a method proposed by Potthoff and Whittinghill, we found significant temporal variation by date of diagnosis at the level of aggregation of one year. However, there was no evidence for general irregular (non-seasonal) temporal clustering within periods less than a year. These results provide little support for the involvement of agents occurring in geographically widespread mini-epidemics, but – taken together with studies of spatial and spatio-temporal clustering - do not preclude the role of more localised sporadic mini-epidemics. Future research should seek to elicit putative environmental agents.

1. INTRODUCTION

Whilst the aetiology of primary biliary cirrhosis (PBC) is not well established, it is likely that both genetics and environmental agents are involved (Kaplan and Gershwin 2005, Hirschfield, Liu et al. 2009, Selmi and Gershwin 2009, Prince, Ducker et al. 2010). An earlier analysis of population-based data on PBC in northeast England showed highly statistically significant space-time clustering, which was most marked for cases diagnosed within 1-4 months of one another (McNally, Ducker et al. 2009). This was the first study to find space-time clustering of this disease and suggests that transient environmental agents might play a role in the aetiology of this disease. Further analysis of this population-based cohort found marked seasonal variation in diagnoses of PBC, with a peak in June (McNally, James et al. 2011). This suggests the involvement of a seasonally varying agent occurring at different locations in the environment. Possible exposures that may be implicated include localised infections associated with agents such as *E. coli*, mycobacteria, and a retrovirus (Butler, Valle et al. 1993, Haydon and Neuberger 2000, Selmi, Balkwill et al. 2003, Xu, Shen et al. 2003). However, the potential role of widespread environmental agents in the aetiology of PBC is unclear. A more recent analysis has demonstrated geographical heterogeneity in PBC in northeast England (McNally, James et al. 2014).

The present study is concerned with the detection of irregular temporal distributions of cases of PBC, as opposed to regular seasonal patterns or space-time clustering as addressed previously (McNally, Ducker et al. 2009, McNally, James et al. 2011), or individual temporal clusters. A general irregular temporal distribution of cases that is not confined to one particular time period or period of each year is known as ‘temporal clustering’ (Muirhead, Cheetham et al. 2013). This sort of clustering could arise because there are a small number of time periods with greatly increased incidence or a large number of time periods with moderately increased incidence. If there were a tendency for cases to

arise in “waves”, then this might indicate that the disease is linked to one or more agents that occur in geographically widespread mini-epidemics; for example, certain widespread infections such as influenza.

2. MATERIALS AND METHODS

2.1 Cases

The data related to incident cases of PBC diagnosed between 1 January 1987 and 31 December 2003 among persons who were resident in a geographically-defined area of northeast England (Northumberland, Sunderland, North Durham, South Durham, Newcastle upon Tyne, North Tyneside, South Tyneside and Gateshead), defined by postcode. Figure 1 shows the study area. The total population of the area at the 2001 census was just over 2 million. The methods of data collection have been described previously (Metcalf, Bhopal et al. 1997). The “date of diagnosis” was defined, as we have previously delineated, as the earliest date at which the patient was found (by examination of clinical case records – hospital or primary care) to have fulfilled any two of the following diagnostic criteria: anti-mitochondrial antibodies (AMA) positive titre ≥ 1 in 40, cholestatic liver blood tests, diagnostic or compatible liver histology (James, Bhopal et al. 1999). The date of diagnosis was determined following examination by the investigators of clinical records and depended upon the date at which the above diagnostic criteria were first fulfilled, rather than being the date at which a diagnosis of PBC was first made and entered in an individual’s clinical case records by the attending doctors.

FIGURE 1 HERE

2.2 Prior Hypothesis

The following aetiological hypothesis was tested: A primary factor influencing temporal heterogeneity of PBC is related to exposure to a geographically widespread, irregularly temporally varying environmental agent occurring either close to diagnosis or at similar times before diagnosis.

2.3 Statistical Methods

The methods described by Muirhead et al (Muirhead, Cheetham et al. 2013) were used to look for clustering in disease rates defined by time periods. This approach is based on a test derived by Potthoff and Whittinghill (Potthoff and Whittinghill 1966, Potthoff and Whittinghill 1966) to look for extra-Poisson variation and assumes that the number of cases in each time period is distributed as negative binomial with the ratio of the variance to the mean equal to a constant, i.e. $1+\beta$. If β is greater than 0 then the observations exhibit extra-Poisson variation, whereas if β equals 0 then the observations are distributed as Poisson. Having conditioned on the total number of cases, the Potthoff-Whittinghill (P-W) statistic is equivalent to:

$$\sum (number\ of\ pairs\ of\ cases\ in\ the\ time\ period / expected\ number\ of\ cases\ in\ the\ time\ period)$$

where the sum is over all the time periods under study. The standardised version of the P-W statistic used here (Muirhead, Cheetham et al. 2013) provides an estimate of β (i.e. the degree of extra-Poisson variation) and is the most powerful test to detect small values of β (Potthoff and Whittinghill 1966, Potthoff and Whittinghill 1966).

In order to distinguish between clustering over the short term and the longer term, a hierarchy of time periods has been considered (Muirhead 2006, Muirhead, Cheetham et al. 2013). For example, the extra-Poisson variation between months within years was estimated by calculating a version of the P-W statistic using the numbers of cases per month within each year and summing this statistic over years. This approach minimises the influence of long-term variation when trying to identify extra-Poisson variation at the level of months. Data aggregated over quarters of a year (defined as January to March; April to June; July to September; and October to December) and over fortnights (defined pragmatically as the first 15 days of the calendar month, or the first 14 days for February, versus the remainder of the month) have also been analysed, in order to look for extra-Poisson variation between fortnights within months, between months within quarters, between quarters within years, etc. Longer-term variation was studied further by testing for extra-Poisson variation in numbers of cases between periods of roughly four years (i.e. 1987-1990, 1991-1994, 1995-1998, 1999-2003).

For analyses of variation between years, the expected numbers of cases took account of year-on-year differences in population sizes; specifically, at ages of 40 years and more, given that PBC is rare among younger people. The expected numbers were standardised on the basis that their total would equal the total number of cases observed. For analyses within any calendar year, the expected number of cases was assumed to be proportional to the length of the period in question, in order to allow for differences in the length of calendar months.

P-values were calculated by conducting 10000 simulations of the standardised P-W statistic under the assumption of Poisson variation, i.e. β equal to 0. Since the focus was on testing for over-dispersion (i.e. $\beta > 0$, reflecting a tendency for cases to aggregate) rather than under-dispersion ($\beta < 0$), one-sided tests were used. In order to examine the robustness of the findings, tests described by Muirhead (Muirhead 2006) were used to see if particular time

periods had a strong influence on the evidence for extra-Poisson variation. Temporal scan statistics were also used to identify periods with particularly high rates, using the SaTScan software¹ (Kulldorff, Heffernan et al. 2005). Also, a periodogram was used to estimate the degree of any periodicity in logs of the rates. Furthermore, because the data were collected in two exercises - but using the same methodology over each time period - roughly before and after the mid-1990s, sensitivity analyses were conducted for diagnoses over the periods 1987-1994 and 1995-2003 separately. Additional subsidiary analyses looked at males and females separately and at diagnoses made at ages <65 y and 65 years or more (NB. there were roughly equal numbers in these two age groups).

3. RESULTS

There were 1029 cases of PBC diagnosed in the study area during 1987-2003 inclusive and whose details were extracted from the population-based register. This number differs slightly from that in previous analyses of these data (McNally, Ducker et al. 2009, McNally, James et al. 2011), owing to the need for full date of diagnosis in the current analysis. Of these patients, 930 were females and 99 were males.

FIGURE 2 HERE

Figure 2 shows the ratio of the observed to the expected number of cases by year of diagnosis. There are indications of periodicity in these data; specifically, a periodogram analysis of the logged annual rates indicates a periodicity of around 9 years. The P-W analysis found strong evidence of extra-Poisson variation between years in the numbers of cases (estimated $\beta = 1.119$, $SE = 0.354$, $P = 0.007$; see the bottom right-hand corner of Table

¹ SaTScanTM is a trademark of Martin Kulldorff. The SaTScanTM software was developed under the joint auspices of (i) Martin Kulldorff, (ii) the National Cancer Institute, and (iii) Farzad Mostashari of the New York City Department of Health and Mental Hygiene. <http://www.satscan.org/>

1). The final year of the study period (i.e. 2003) - during which 90 cases were diagnosed compared with an expected value of 63.4 – had a high influence on this result; excluding this year from the analysis changed the P-value to 0.091. A temporal scan analysis also identified this year as having a particularly high rate. However, a sensitivity analysis provided strong evidence ($P=0.011$) of extra-Poisson variation between periods of roughly four years, so confirming this long-term variation in PBC cases; this evidence persisted even after excluding data for 2003 ($P=0.005$). There was also evidence of variation between four-year periods during 1997-1994 and between years during 1995-2003 (see Table 2). The evidence for extra-Poisson variation between years was similar for diagnoses at ages <65 and at 65 years or more ($P=0.034$ and 0.013 respectively), but was stronger among males than among females ($P=0.001$ and $P=0.13$ respectively). Nevertheless, the standardised P-W statistic for the analysis between years remained highly significant when averaged either over ages or over sexes ($P=0.004$ and $P=0.001$ respectively), so indicating that the analysis between years has not been confounded by these factors.

TABLE 1 HERE

In contrast to the analyses between years, none of the P-W analyses either within months, within quarters of a year or within years found statistically significant evidence of extra-Poisson variation in the numbers of cases of PBC (see Table 1, excluding the final column). Repeating these analyses for diagnoses during 1987-1994 and 1995-2003 separately generally gave similar results (see Table 2); the only exception was evidence for extra-Poisson variation between months within quarters over 1987-1994 ($P=0.029$) but not over 1995-2003 ($P=0.50$). In this regard, it can be seen from Figure 3 that the June peak in PBC by month of diagnosis reported previously (McNally, James et al. 2011) is particularly pronounced in the earlier period, indicating some impact of this seasonal variation on the P-W analysis between months within quarters. Analyses conducted separately for males and

females and separately for those aged less than 65 years and those aged 65 and over did not show extra-Poisson variation between periods less than a year (results not shown), with the exception of the analysis between months within quarters at ages 65 years and above ($P=0.048$) and the analysis between fortnights within quarters for males ($P=0.035$).

TABLE 2 HERE

FIGURE 3 HERE

4. DISCUSSION

This analysis is the first of its type to study temporal clustering of this autoimmune liver disease. The methods that have been used are statistically rigorous and have been applied to analyse high-quality, population-based data from a well-defined geographical region. We identified long-term temporal variation in diagnoses of PBC, but found little evidence of short-term (non-seasonal) variation in the numbers of cases. This pattern of occurrence is not consistent with the involvement of an agent that displays non-regular temporal epidemicity. For example, the findings are not consistent with an infection that occurs in geographically widespread mini-epidemics, e.g. influenza. However, the results do not exclude the possible role of localised infectious agents, such as mycobacteria.

It might be queried whether these findings are artefacts, due to the way in which cases were diagnosed; for example, whether one patient's diagnosis might provoke a general practitioner to consider the diagnosis of PBC in another patient, or whether any batching of assessments of AMAs might lead to “bunching” of cases. However, neither of these is at all likely. Firstly, the diagnostic criteria described earlier are robust and well-defined. Secondly, the timescale between taking samples and reporting AMA findings is a matter of days, rather

than weeks or months. Furthermore, the date of diagnosis often precedes the onset of symptoms and the first appointment with a specialist.

The long-term temporal variation apparent from Figure 2 indicates that some environmental factor might be involved in the aetiology of PBC. Although the statistical significance is affected particularly by the large number of cases observed in the last year of the study period, there is little reason to think that this long-term pattern is an artefact due to variations in data collection. In particular, a sensitivity analysis that looked separately at the two data collection periods found evidence of long-term variation over both 1987-1994 and 1995-2003. Another possibility is that changes over time in the size and/or age and sex structure of the study population might have affected the annual numbers of cases. However, the study area has markedly low inward and outward migration rates (Office of Population Censuses and Surveys Census Division and General Registrar Office (Scotland) Census Branch 1983, Office of Population Censuses and Surveys Census Division and General Registrar Office (Scotland) Census Branch 1991, Office of Population Censuses and Surveys Census Division, General Registrar Office (Scotland) Census Branch et al. 2001) and population sizes were generally stable over the period studied. Furthermore, adjustments for age and sex did not reduce the statistical significance of the variation between years.

Table 3 summarises key findings from studies in northeast England of spatial, spatio-temporal and temporal clustering of PBC, as well as of seasonal variation, following the approach taken in an earlier assessment of clustering of childhood cancer (McNally and Eden 2004). It is very important to note that the lack of temporal clustering seen here does not contradict our earlier analyses (McNally, Ducker et al. 2009, McNally, James et al. 2011, McNally, James et al. 2014). The finding of spatial clustering indicates the involvement of a localised environmental factor (McNally, James et al. 2014), whilst the space-time clustering is consistent with the involvement of an aetiological agent that has an irregular occurrence in

time and space, taking into account general excesses in certain geographical areas or temporal periods (McNally, Ducker et al. 2009). Also, the finding of seasonal variation is consistent with the involvement of a regularly occurring seasonal exposure in aetiology (McNally, James et al. 2011). In contrast to the latter analysis, the current analysis did not examine whether PBC diagnoses tend to occur more often than expected in the same month or same season each year. Rather, we examined here whether diagnoses tend to cluster together in time, irrespective of any periodicity in this pattern. By its design, the earlier analysis (McNally, James et al. 2011) had greater statistical power to detect seasonal variation, whereas the current analysis has greater power to detect a wider range of temporal variations, including irregular patterns. Furthermore, the seasonal pattern reported previously (McNally, James et al. 2011) appears to have had only a marginal impact on the current analysis.

Our present study has indicated that PBC does not occur in “mini-epidemics” that are geographically widespread and occur at specific points in time. This suggests that, for example, widespread infections are not involved in the aetiology of PBC. However, the results do not preclude the involvement of more localised sporadic “mini-epidemics” that lead to space-time clustering, nor do they exclude seasonality. Thus, a range of putative agents could be involved, including infectious and non-infectious agents, as discussed further by (McNally, Ducker et al. 2009, McNally, James et al. 2011, McNally, James et al. 2014). Conceivably exposure to different putative agents, occurring in the environment from time to time, might be involved in stimulating the immune process responsible for PBC in susceptible individuals, each responsible for a mini-epidemic. A similar mechanism has been proposed for childhood precursor B-cell acute lymphoblastic leukaemia (Greaves 1988). Possible infectious agents, indicated by other studies, include *E. coli*, *Novosphingobium aromaticovorans*, mycobacteria and human beta retrovirus (Butler, Valle et al. 1993, Butler, Hamilton-Miller et al. 1995, Haydon and Neuberger 2000, Shimoda, Nakamura et al. 2000,

Tsuneyama, Harada et al. 2001, Selmi, Balkwill et al. 2003, Xu, Shen et al. 2003, Kaplan 2004, Olafsson, Gudjonsson et al. 2004, Gershwin, Selmi et al. 2005, Padgett, Selmi et al. 2005, Mattner, Savage et al. 2008). Possible non-infectious agents include xenobiotics that commonly occur in pollutants, food preservatives and pesticides (Dronamraju, Odin et al. 2010), as well as past or present smoking and previous pregnancies (Selmi, Bowlus et al. 2011). Whilst some of these factors suggest a long latency for the onset of clinical disease, the potential role of infectious agents – together with the seasonal pattern reported previously (McNally, James et al. 2011) – suggest that the latency may be short for certain risk factors.

In conclusion, this study has not found any evidence for general irregular (non-seasonal) temporal clustering among cases of PBC. This finding is highly novel in itself. It suggests that there is little support for the involvement of agents that occur in geographically widespread mini-epidemics; for example, certain widespread infections such as influenza, in contrast to more localised infectious agents such as mycobacteria. Future research should seek to elicit other putative environmental agents (infectious or non-infectious) that might explain the long-term temporal variation seen for PBC.

Funding: This work was supported by the Medical Research Council (UK) and the Wellcome Trust. The sponsors had no involvement in any aspect of this study.

Conflicts of Interest: None.

Ethics approval: Case selection was approved by local ethical committees.

REFERENCES

- Butler, P., J. Hamilton-Miller, H. Baum and A. K. Burroughs (1995). "Detection of M2 antibodies in patients with recurrent urinary tract infection using an ELISA and purified PBC specific antigens. Evidence for a molecular mimicry mechanism in the pathogenesis of primary biliary cirrhosis?" Biochem Mol Biol Int **35**(3): 473-485.
- Butler, P., F. Valle, J. M. Hamilton-Miller, W. Brumfitt, H. Baum and A. K. Burroughs (1993). "M2 mitochondrial antibodies and urinary rough mutant bacteria in patients with primary biliary cirrhosis and in patients with recurrent bacteriuria." J Hepatol **17**(3): 408-414.
- Dronamraju, D., J. Odin and N. Bach (2010). "Primary biliary cirrhosis: environmental risk factors." Dis Markers **29**(6): 323-328.
- Gershwin, M. E., C. Selmi, H. J. Worman, E. B. Gold, M. Watnik, J. Utts, K. D. Lindor, M. M. Kaplan and J. M. Vierling (2005). "Risk factors and comorbidities in primary biliary cirrhosis: a controlled interview-based study of 1032 patients." Hepatology **42**(5): 1194-1202.
- Greaves, M. F. (1988). "Speculations on the cause of childhood acute lymphoblastic leukemia." Leukemia **2**(2): 120-125.
- Haydon, G. H. and J. Neuberger (2000). "PBC: an infectious disease?" Gut **47**(4): 586-588.
- Hirschfield, G. M., X. Liu, C. Xu, Y. Lu, G. Xie, Y. Lu, X. Gu, E. J. Walker, K. Jing, B. D. Juran, A. L. Mason, R. P. Myers, K. M. Peltekian, C. N. Ghent, C. Coltescu, E. J. Atkinson, E. J. Heathcote, K. N. Lazaridis, C. I. Amos and K. A. Siminovitch (2009). "Primary biliary cirrhosis associated with HLA, IL12A, and IL12RB2 variants." N Engl J Med **360**(24): 2544-2555.
- James, O. F., R. Bhopal, D. Howel, J. Gray, A. D. Burt and J. V. Metcalf (1999). "Primary biliary cirrhosis once rare, now common in the United Kingdom?" Hepatology **30**(2): 390-394.
- Kaplan, M. M. (2004). "Novosphingobium aromaticivorans: a potential initiator of primary biliary cirrhosis." Am J Gastroenterol **99**(11): 2147-2149.
- Kaplan, M. M. and M. E. Gershwin (2005). "Primary biliary cirrhosis." N Engl J Med **353**(12): 1261-1273.
- Kulldorff, M., R. Heffernan, J. Hartman, R. Assuncao and F. Mostashari (2005). "A space-time permutation scan statistic for disease outbreak detection." PLoS Med **2**(3): e59.
- Mattner, J., P. B. Savage, P. Leung, S. S. Oertelt, V. Wang, O. Trivedi, S. T. Scanlon, K. Pendem, L. Teyton, J. Hart, W. M. Ridgway, L. S. Wicker, M. E. Gershwin and A. Bendelac (2008). "Liver autoimmunity triggered by microbial activation of natural killer T cells." Cell Host Microbe **3**(5): 304-315.
- McNally, R. J. and T. O. Eden (2004). "An infectious aetiology for childhood acute leukaemia: a review of the evidence." Br J Haematol **127**(3): 243-263.

McNally, R. J., P. W. James, S. Ducker, P. D. Norman and O. F. James (2014). "No rise in incidence but geographical heterogeneity in the occurrence of primary biliary cirrhosis in North East England." Am J Epidemiol **179**(4): 492-498.

McNally, R. J. Q., S. Ducker and O. F. James (2009). "Are transient environmental agents involved in the cause of primary biliary cirrhosis? Evidence from space-time clustering analysis." Hepatology **50**(4): 1169-1174.

McNally, R. J. Q., P. W. James, S. Ducker and O. F. James (2011). "Seasonal variation in the patient diagnosis of primary biliary cirrhosis: Further evidence for an environmental component to etiology." Hepatology **54**(6): 2099-2103.

Metcalf, J. V., R. S. Bhopal, J. Gray, D. Howel and O. F. James (1997). "Incidence and prevalence of primary biliary cirrhosis in the city of Newcastle upon Tyne, England." Int J Epidemiol **26**(4): 830-836.

Muirhead, C. R. (2006). "Methods for detecting disease clustering, with consideration of childhood leukaemia." Stat Methods Med Res **15**(4): 363-383.

Muirhead, C. R., T. D. Cheetham, S. Court, M. Begon and R. J. McNally (2013). "How do childhood diagnoses of type 1 diabetes cluster in time?" PLoS One **8**(4): e60489.

Office of Population Censuses and Surveys Census Division and General Registrar Office (Scotland) Census Branch (1983). 1981 census: small area statistics [computer file]. Colchester, Essex, UK, UK Data Service

Office of Population Censuses and Surveys Census Division and General Registrar Office (Scotland) Census Branch (1991). 1991 census: small area statistics and local base statistics [computer file]. Colchester, Essex, UK, UK Data Service

Office of Population Censuses and Surveys Census Division, General Registrar Office (Scotland) Census Branch and Northern Ireland Statistics and Research Agency (2001). 2001 census: aggregate statistics [computer file]. Colchester, Essex, UK, UK Data Service

Olafsson, S., H. Gudjonsson, C. Selmi, K. Amano, P. Invernizzi, M. Podda and M. E. Gershwin (2004). "Antimitochondrial antibodies and reactivity to N. aromaticivorans proteins in Icelandic patients with primary biliary cirrhosis and their relatives." Am J Gastroenterol **99**(11): 2143-2146.

Padgett, K. A., C. Selmi, T. P. Kenny, P. S. Leung, D. L. Balkwill, A. A. Ansari, R. L. Coppel and M. E. Gershwin (2005). "Phylogenetic and immunological definition of four lipoylated proteins from *Novosphingobium aromaticivorans*, implications for primary biliary cirrhosis." J Autoimmun **24**(3): 209-219.

Potthoff, R. F. and M. Whittinghill (1966). "Testing for homogeneity. I. The binomial and multinomial distributions." Biometrika **53**(1): 167-182.

Potthoff, R. F. and M. Whittinghill (1966). "Testing for homogeneity. II. The Poisson distribution." Biometrika **53**(1): 183-190.

Prince, M. I., S. J. Ducker and O. F. James (2010). "Case-control studies of risk factors for primary biliary cirrhosis in two United Kingdom populations." Gut **59**(4): 508-512.

Selmi, C., D. L. Balkwill, P. Invernizzi, A. A. Ansari, R. L. Coppel, M. Podda, P. S. Leung, T. P. Kenny, J. Van De Water, M. H. Nantz, M. J. Kurth and M. E. Gershwin (2003). "Patients with primary biliary cirrhosis react against a ubiquitous xenobiotic-metabolizing bacterium." Hepatology **38**(5): 1250-1257.

Selmi, C., C. L. Bowlus, M. E. Gershwin and R. L. Coppel (2011). "Primary biliary cirrhosis." Lancet **377**(9777): 1600-1609.

Selmi, C. and M. E. Gershwin (2009). "The role of environmental factors in primary biliary cirrhosis." Trends Immunol **30**(8): 415-420.

Shimoda, S., M. Nakamura, H. Shigematsu, H. Tanimoto, T. Gushima, M. E. Gershwin and H. Ishibashi (2000). "Mimicry peptides of human PDC-E2 163-176 peptide, the immunodominant T-cell epitope of primary biliary cirrhosis." Hepatology **31**(6): 1212-1216.

Tsuneyama, K., K. Harada, N. Kono, K. Hiramatsu, Y. Zen, Y. Sudo, M. E. Gershwin, M. Ikemoto, H. Arai and Y. Nakanuma (2001). "Scavenger cells with gram-positive bacterial lipoteichoic acid infiltrate around the damaged interlobular bile ducts of primary biliary cirrhosis." J Hepatol **35**(2): 156-163.

Xu, L., Z. Shen, L. Guo, B. Fodera, A. Keogh, R. Joplin, B. O'Donnell, J. Aitken, W. Carman, J. Neuberger and A. Mason (2003). "Does a betaretrovirus infection trigger primary biliary cirrhosis?" Proc Natl Acad Sci U S A **100**(14): 8454-8459.

Table 1: Application of the Potthoff-Whittinghill technique to detect temporal clustering of PBC cases in northeast England during 1987-2003 inclusive

Type of analysis	PW (SE) <i>one-sided P-value</i>			
	Within months	Within quarters	Within years	Within full study period
Between fortnights	-0.126 (0.116) <i>P=0.87</i>	0.012 (0.080) <i>P=0.44</i>	-0.019 (0.072) <i>P=0.59</i>	0.030 (0.070) <i>P=0.32</i>
Between months		0.175 (0.126) <i>P=0.087</i>	0.042 (0.104) <i>P=0.33</i>	0.119 (0.099) <i>P=0.12</i>
Between quarters			-0.226 (0.200) <i>P=0.88</i>	0.096 (0.173) <i>P=0.28</i>
Between years				1.119 (0.354) <i>P=0.007</i>

Notes:

- PW is the one-step estimate of β , the extra-Poisson variation, calculated as $S/i(0)$ in the notation of Muirhead (2006).
- SE is the standard error of PW in the absence of extra-Poisson variation, calculated as $1/\sqrt{i(0)}$ in the notation of Muirhead (2006).
- P-values have been calculated using 10000 simulations of PW, assuming Poisson variation. All P-values are one-sided.

Table 2: Application of the Potthoff-Whittinghill technique to detect temporal clustering of PBC cases in northeast England, conducted separately for 1987-1994 and 1995-2003.

Type of analysis	Time period	PW (SE) <i>one-sided p-value</i>				
		Within months	Within quarters	Within years	Within four-year periods	Within full study period
Between fortnights	1987-1994	-0.114 (0.169) <i>P=0.76</i>	0.090 (0.116) <i>P=0.22</i>	0.053 (0.105) <i>P=0.29</i>	0.025 (0.103) <i>P=0.39</i>	0.074 (0.102) <i>P=0.23</i>
	1995-2003	-0.136 (0.159) <i>P=0.81</i>	-0.058 (0.110) <i>P=0.70</i>	-0.083 (0.099) <i>P=0.79</i>	-0.025 (0.097) <i>P=0.59</i>	-0.018 (0.097) <i>P=0.56</i>
Between months	1987-1994		0.379 (0.183) <i>P=0.029</i>	0.198 (0.152) <i>P=0.10</i>	0.129 (0.146) <i>P=0.18</i>	0.193 (0.145) <i>P=0.097</i>
	1995-2003		-0.007 (0.173) <i>P=0.50</i>	-0.096 (0.143) <i>P=0.74</i>	0.017 (0.138) <i>P=0.43</i>	0.036 (0.137) <i>P=0.37</i>
Between quarters	1987-1994			-0.187 (0.291) <i>P=0.72</i>	-0.256 (0.259) <i>P=0.85</i>	-0.069 (0.254) <i>P=0.58</i>
	1995-2003			-0.262 (0.275) <i>P=0.83</i>	0.170 (0.243) <i>P=0.23</i>	0.224 (0.239) <i>P=0.17</i>
Between years	1987-1994				-0.612 (0.579) <i>P=0.89</i>	0.262 (0.535) <i>P=0.27</i>
	1995-2003				1.923 (0.536) <i>P=0.004</i>	1.937 (0.500) <i>P=0.003</i>
Between four-year periods	1987-1994					5.437 (1.416) <i>P=0.010</i>
	1995-2003					1.278 (1.416) <i>P=0.13</i>

Notes:

- PW is the one-step estimate of β , the extra-Poisson variation, calculated as $S/i(0)$ in the notation of Muirhead (2006).

- SE is the standard error of PW in the absence of extra-Poisson variation, calculated as $1/\sqrt{i(0)}$ in the notation of Muirhead (2006).
- P-values have been calculated using 10000 simulations of PW, assuming Poisson variation. All P-values are one-sided.
- The “four-year periods” are 1987-1990, 1991-1994, 1995-1998 and 1999-2003. Allowance has been made for differences in the length of these periods.

Table 3: Findings from analyses in northeast England of spatial, spatio-temporal and temporal clustering of PBC, as well as of seasonal variation

Type of analysis (Reference)	Spatial clustering (McNally, James et al. 2014)	Spatio-temporal clustering (McNally, Ducker et al. 2009)	Temporal clustering (This paper)	Seasonal variation (McNally, James et al. 2011)
Key findings	Point process methods indicated presence of spatial clustering ($P = 0.001$).	Analysis using Kulldorff's scan statistic showed highly statistically significant space-time clustering ($P < 0.001$). Clustering was most marked for cases diagnosed within 1-4 months of one another.	Using an adapted version of the Potthoff-Whittinghill method, there was significant temporal variation by date of diagnosis at the level of aggregation of one year, but no evidence for general irregular (non-seasonal) temporal clustering within periods less than a year.	Poisson regression analysis provided evidence of a sinusoidal pattern with a June peak ($P = 0.012$).
Implications (<i>after (McNally and Eden 2004)</i>)	This may indicate the involvement of a localised environmental factor, such as an infection.	This suggests a role for transient environmental agents, such as an infection occurring in mini-epidemics.	This does not support the involvement of agents occurring in geographically widespread mini-epidemics, but a role for more localised sporadic mini-epidemics cannot be excluded.	This provides evidence for the involvement of a seasonally varying environmental agent in the aetiology of PBC.

Figure 1: Map of the study area in northeast England

Figure 2: Ratio of observed to expected number of cases of PBC in northeast England by year of diagnosis

Figure 3: Number of cases of PBC in northeast England by calendar month and period of diagnosis