The latitude gradient for multiple sclerosis prevalence is established in the early lifecourse

Clive E. Sabel,[1,†] John F. Pearson,[2,†] Deborah F. Mason,[3] Ernest Willoughby,[4] David A. Abernethy[5] and Bruce V. Taylor[6]

***These authors contributed equally to this work.**

Abstract

The strongest epidemiological clue that the environment at the population level has a significant impact on the risk of developing multiple sclerosis (MS) is the well-established, and in many instances, increasing latitudinal gradient of prevalence, incidence and mortality globally, with prevalence increasing by up to 10-fold between the equator and 60 degrees North and South. The drivers of this gradient are thought to be environmental with latitude seen as a proxy for ultraviolet radiation and thus vitamin D production, however other factors may also play a role. However several important questions remain unanswered, particularly when in the life course is the gradient established, does lifetime migration mitigate or exacerbate previously reported latitude gradients at location of diagnosis, and do factors such as sex or MS disease phenotype influence the timing or significance of the gradient?

Utilising life time residence calendars collected as part of the New Zealand national MS prevalence study, we constructed lifetime latitudinal gradients for MS from birth to prevalence day 2006 taking into account migration internally and externally and then analysed by sex and MS clinical course phenotype. 2127 of 2917 people living in NZ on prevalence day 7 March 2006 with MS completed the life course questionnaire and of these 1587 were born in NZ. All cohorts and sub cohorts were representative of the overall MS population in NZ on prevalence day.

We found that the prevalence gradient was present at birth and was in fact stronger than at census day, and the slope of the gradient persisted until the age of 12 before gradually declining. We found that internal and external migration into NZ had little if any effect on the gradient except to decrease the significance of the gradient somewhat. Finally, we found as we had reported previously that the lifetime prevalence gradients were largely driven by females with relapse onset MS.

These findings confirm for the first time the importance of early life environmental exposures in the risk of MS indicating strongly that exposures as early as in utero and at birth drive the latitudinal

gradient. Consequently, prevention studies should be focussed on high risk individuals and populations from the earliest possible time points especially, when appropriate, on females.

Author affiliations:

1 BERTHA, Big Data Centre for Environment and Health, Department of Environmental Science, Aarhus University, Denmark

2 Biostatistics and Computational Biology Unit, University of Otago, Christchurch, New Zealand

3 Christchurch Hospital, Christchurch, New Zealand

4 Auckland City Hospital, Auckland, New Zealand

5 Capital and Coast Hospital, Wellington, New Zealand

6 Menzies Institute for Medical Research, University of Tasmania, Australia

Correspondence to: Clive Sabel

BERTHA, Big Data Centre for Environment and Health, Department of Environmental Science, Aarhus University, Denmark

E-mail: cs@envs.au.dk

Running head: Early life established MS latitude gradient

Keywords: multiple sclerosis; New Zealand; latitude gradient; early life exposome

1. Introduction

Numerous studies, over many decades have reported distinct geographical and temporal patterns of multiple sclerosis prevalence, incidence, and mortality. Broadly, in both the northern and southern hemispheres, increasing latitude towards the poles, is associated with increasing prevalence (Kurtzke, 1977, 2000; Mayer, 1981; Ebers and Sadovnick, 1993; Wallin *et al.*, 2004; Taylor *et al.*, 2010a; Risco

et al., 2011; Simpson *et al.*, 2011, 2019). Similarly, there has been a clear increase in multiple sclerosis prevalence over time.

It has been hypothesised that latitude is a surrogate for an environmental exposure, such as UV exposure, and therefore by inference Vitamin D levels, as up to 95% of vitamin D in most societies is generated by the exposure of precursors in the skin to ultraviolet B (UVB) radiation (Lucas, 2013; Simpson, 2018). However, although there is significant epidemiological evidence for a link between multiple sclerosis risk and lower UVB/vitamin D there is little evidence that vitamin D supplementation or increased UVB decreases an individual's or a populations risk of developing multiple sclerosis. This is largely due to the long latency or prodromal period before multiple sclerosis becomes symptomatic and the overall low prevalence of the disease, which makes intervention studies of this type impractical.

Importantly, other factors including different rates and timing of Epstein Barr Virus (EBV) infection, ethnic/genetic make-up and even magnetic field fluctuations have also been suggested as contributing to the latitudinal gradient of multiple sclerosis prevalence (Simpson *et al.*, 2019). It is also clear that multiple sclerosis risk is mediated by other personal, environmental, and genetic factors and that the weighting of these factors in multiple sclerosis risk may have significantly differential effects between individuals and populations to reach a causal sufficiency (Rothman's causal pie, Rothman, 1976). Multiple sclerosis risk has been associated with genetic factors, explaining up to 25% of the risk (IMSGC, 2019), EBV exposure and timing and response to EBV infection (particularly the development of infectious mononucleosis), cigarette smoking, adolescent obesity and other factors. In epidemiological studies of those presenting with a first episode of multiple sclerosis, combinations of these factors have been shown to explain between 60% and 80% of multiple sclerosis risk (van der Mei *et al.*, 2016). However, where you live remains one of the most important drivers of multiple sclerosis risk, with the latitudinal gradient persisting and even strengthening over time (Simpson, 2019).

Many studies examining associations between geographical patterns of disease and disease causation assume that current residence can be equated with exposure to conditions that currently (and historically) pertain at that location (Bentham, 1988). This is important, since the place of residence at the time of diagnosis or death is often adopted by epidemiologists and geographers as the location for further analysis of the disease in question. However, people move throughout their lifetimes, and hence previous exposure to environmental factors at other locations, will not typically be included in any study. By adopting only the current residential address, not only will an individual's long-term migration history be neglected, but additionally the daily 'activity spaces' of cases will be also be

ignored (Sabel *et al.*, 2000). In short, study designs that look beyond the place of diagnosis (Sabel *et al.*, 2003) should be important to unpack the latitudinal gradient for multiple sclerosis, and recently the concepts of whole life course (Jacquez *et al.*, 2015), and the exposome (Wild, 2005) have been conceptualised to address this.

The life course approach in spatial epidemiology (Jia *et al.*,, 2019), with its roots in Time Geography (Hägerstrand, 1970) explicitly recognises that individual mobility across the whole life course from (pre-) birth to diagnosis, is a useful construct when considering diseases with a long latency period between exposure and diagnosis. The concept of the exposome was introduced by Wild (2005) to describe accumulated life course environmental exposures (including lifestyle factors), from the prenatal period onwards. The exposome measures the effects of lifelong exposure to environmental influences on human health and therefore requires a longitudinal study design. The latitude effect that we explore in this paper, whether that is Vitamin D, sunlight, or some other environmental exposure that varies by latitude, is one component of the Exposome.

Concepts such as the life course and the exposome are particularly pertinent when studying multiple sclerosis where the exact timing of exposures that may relate to the onset of the disease is unknown. Exposure during pregnancy may well play an important role in the subsequent development of multiple sclerosis with month of birth studies from both the Northern and Southern hemispheres demonstrating an effect of gestation in winter as predicting subsequent risk of multiple sclerosis (Dobson *et al.*, 2013). Similarly studies have demonstrated parity (Ponsonby, 2012) and events in teenage years (EBV infection leading to Infectious mononucleosis (Lucas *et al.*, 2011a)), early adulthood smoking (Pakpoor *et al.*, 2020), and obesity (Pakpoor *et al.*, 2020) may affect the subsequent risk of multiple sclerosis. Therefore, it is important to know when the strongest indicator of the environmental effects on multiple sclerosis onset (the latitudinal gradient of multiple sclerosis prevalence) is established.

The timing of establishment of a person's risk of multiple sclerosis may be very important for development of strategies for its prevention. For example, if the gradient is driven by Ultraviolet Radiation (UVR) from sun exposure in the years prior to the clinical manifestations of multiple sclerosis then appropriate guidance on sun exposure/vitamin D supplementation can be directed at young adults. If however the gradient is driven by environmental factors at or around birth then the same messages need to be directed towards pregnant women.

Population migration constrains the validity of inferences from prevalence models of multiple sclerosis, particularly when inferring disease aetiology. Most multiple sclerosis studies which have looked at migration have concentrated on the effect of age at migration in conjunction with the

original and final residence locations (Alter *et al.*, 1978; Elian *et al.*, 1990; Wallin *et al.*, 2009; McLeod *et al.*, 2011). However, Cabre *et al.* (2005) also considered the effect of migrants returning to the French West Indies from countries such as France, where the prevalence of multiple sclerosis is much higher. For that study, where 33% of the population were migrants, the prevalence of multiple sclerosis was higher in the migrants than in the non-migrants by a factor of 2.2. In studies, where migrants are moving to an area of high prevalence, the migrant groups tend to have a lower rate of prevalence than the resident group, but second-generation migrants more closely resemble the resident group (Munk Nielsen, 2019). In most multiple sclerosis migration studies, the presence of migrants within the wider population cohort affects the prevalence of multiple sclerosis within that population (Munk Nielsen, 2019). To understand the aetiology of multiple sclerosis, migration effects must be considered when determining the prevalence. In addition to allowing for external migration, care must be taken to control for internal migration if effects which could vary with migration are to be analysed. Prevalence – latitude studies which do not consider lifetime internal migration may therefore be allowing migration to influence their results.

The most significant finding from Taylor *et al.* (2010a) was the unequivocal confirmation of a highly significant latitudinal gradient of multiple sclerosis prevalence in NZ at the 2006 time point, with prevalence increasing threefold between the north and south of the country. This finding was the basis upon which to explore the impact of lifetime migration on the latitudinal gradient over time in this paper. Here, we aim to answer two questions: Does lifetime population migration dilute or concentrate the previously reported latitudinal gradient in multiple sclerosis in the NZ multiple sclerosis prevalence (NZMSP) study; and when in the life course is the effect of latitude on multiple sclerosis risk established?

2. Materials and Methods

2.1 Data

We used data from the New Zealand (NZ) national multiple sclerosis prevalence study (NZNMSPS) conducted on the national population census day, 7th March 2006. This study has been reported elsewhere (Taylor *et al.*, 2010a) but briefly 2917 persons living with multiple sclerosis were identified as resident in NZ on prevalence day. Capture recapture analyses indicated that greater than 96% of all people with multiple sclerosis in NZ on that day were identified. Of these 2917 cases, 2127 completed a comprehensive questionnaire, aided by a research nurse if necessary, of which 1587 were

both born in NZ, and able to complete their full lifetime residential history from conception to diagnosis of multiple sclerosis (hereafter the NZ Born sub cohort). Responders were on average 2 years younger than non-responders but there was no significant difference in sex ratio or distribution of symptoms at onset. The age difference, however, was not significantly different between responders born in New Zealand and the entire cohort (Table 1.) We used a concept of epochs to distinguish between unique residential locations for each case, where an epoch is a time period declared by the respondent, for each permanent residential location at which they have ever resided. It could be as short as 1 day, or as long as a whole lifetime. Of the 1587 sub cohort, 1167 recorded all residential epochs as being within NZ.

We, Taylor *et al.*, (2010a), established a robust latitudinal gradient for multiple sclerosis prevalence at census day 2006 in NZ. Here we used the same base data from the NZ multiple sclerosis prevalence study, but limited to the NZ born population, to estimate the latitudinal gradient of multiple sclerosis from birth, and to compare this gradient throughout the life course, up to prevalence day 2006. New Zealand is ideally suited to such a study as it is latitudinally extensive, ranging from 35S to 48S with a widely dispersed population, largely of Northern European origin. The NZ health system is highly developed with equitable access to neurological care throughout the country.

The geography, and the choice of zoning we adopted, needs careful consideration, and was designed to acknowledge and minimise the challenges posed by the modifiable areal unit problem for health studies (Flowerdew *et al.*, 2008). The census geography of New Zealand has changed considerably over the last century undergoing various revisions, and the population age and sex structure has similarly also changed considerably. Population censuses are available generally every 5 years starting in 1851. Prior to 1989, New Zealand had a system of counties and boroughs. The 1989 local government reform created 13 regions, which formed the basis for our analysis where the 13 NZ census regions in 2006 were aggregated into six broad latitudinal regions from North to South (Fig. 1). These latitudinal regions were constructed so that each contains sufficient multiple sclerosis cases to allow meaningful stratification by ethnicity, age, sex and multiple sclerosis phenotype whilst minimising the potential for latitude misclassification. For each region, a population weighted latitude and longitude centroid (PWC) was calculated, and this centroid was taken as the latitudinal reference point for that region and thus subsequent analysis (Fig. 1).

2.2 Statistical Analysis

Participants who responded to the questionnaire and provided residence data were stratified by region of birth and further stratified by residence in New Zealand until diagnosis. There were 6 people who

did not provide their date of diagnosis and this was imputed at the median age by sex and age group on census day. Sub cohorts were compared with the full cohort using Wilcoxon rank sum tests for year of birth, Fisher exact tests for sex and relapsing vs progressive onset and Kolmogorov-Smirnov tests for location on census day. Age was split into 6 groups (under 25, 25-34, 35-44, 45-54, 55-64 and over 64 years) to allow sufficient people with multiple sclerosis in each group and these age groups were used to calculate weights in the New Zealand usually resident population on census day 7 March 2006. There was no statistical difference observed in birth year, sex, relapsing onset or location at census day between the entire cohort and the 1,587 (54.4%) New Zealand born sub-cohort (excluding foreign born multiple sclerosis cases) used in the subsequent analysis (Table 1). Those born elsewhere were slightly older (2 years for the median) and tended to reside more in the Auckland region (region 1), consistent with NZ immigration patterns.

Prevalence based on location at census day in 2006 for NZ Born people with multiple sclerosis (PwMS) was directly age standardised using the calculated weights, thereby removing any age bias in regions that might be younger than others. These weights were chosen to minimise distortion in confidence intervals induced by weighting to a different population. Latitude gradients were calculated by linear regression of prevalence in regions against the latitude of the population weighted centroid. Gradients between cohorts were compared by testing for statistical significance of the interaction term between cohort (dichotomous) and latitude (continuous) in the linear model of prevalence on cohort and latitude. For convenience, gradients are reported as positive and latitudes are graphed as positive, although as New Zealand is in the southern hemisphere, conventionally latitudes and hence gradients are negative. Prevalence and gradients weighted to standard populations have been provided earlier.

2.3 Data Availability

The data that support the findings of this study are available on reasonable request from the corresponding author. The data are not publicly available due to their containing information that would compromise the privacy of research participants.

3. Results

3.1 Comparing Total Case prevalence with the NZ Born sub-cohort prevalence

To address our first hypothesis; whether lifetime population migration dilutes or concentrates the previously reported latitude gradient of prevalence in multiple sclerosis in the NZMSP study, in Table 2 and Figure 2, we compared the total case cohort with the NZ Born cohort prevalence gradients both at birth and 2006 location. This showed a higher, but not statistically significant (P=0.14) gradient for location of birth when compared to location in 2006. Thus the pattern observed in 2006 is also present in the birthplace location, but with a stronger signal, consistent with the latitudinal gradient being established prior to lifetime migration. What is crucial here is not whether the gradient at birth is significantly higher than in 2006, but that the timing of establishment has been shown to exist at birth, much earlier in the life course than diagnosis or prevalence day in 2006.

3.2 Latitude Gradients by Phenotype and Sex

Our next analysis in Fig. 3 explores whether the established gradients vary by sex and disease course phenotype. The key message is that the process is driven by females with relapsing onset multiple sclerosis (ROMS).

3.3 Latitude Gradient by Age

Having established that the latitude gradient exists, and is higher, but not statistically different, for location at birthplace than for location in 2006, we now address our second question; namely when in the life course is the effect of latitude on multiple sclerosis risk established?

In Fig. 4, we explore the cohort prevalence for the NZ born cohort going back in time. Cases are the NZ born multiple sclerosis cohort, by location at each birthday, taken from the residential epoch data. The denominator is the total NZ born population in 2006 (thus the denominator is constant), with their location at census day 7/3/2006. Our analysis is limited by having epoch data only for the cases. We elaborate in the discussion on the limitation of not knowing the denominator population locations at time points prior to 2006 and thus can account for migration only in cases, but not for migration in the total NZ population. Note that this is not the gradient for a given calendar year, but rather the gradient for cases at each specified age (which will occur in different calendar years, since cases were born from 1913 to 1990).

Focusing on the dark blue line in Fig. 4, we show that the high latitude gradient from south to north was present at birth in those born in NZ and that this gradient remained relatively stable just over 10, until the age of about 12, and then declined as the cases migrate (light grey, at bottom of graph), to dilute the effect. This gradient can be interpreted as the increase in age-adjusted prevalence per

100,000 persons for a degree increase in latitude. Although the latitude gradients at ages earlier than 45 are not statistically different from the 2006 latitude gradient, there is a clear asymptotic pattern below 12 years of age, followed by a drop off in gradient that asymptotes at later ages, although this will be somewhat affected by power. Modelling the curve with an unconstrained 5 parameter logistic regression estimates the asymptotic gradient at birth to be 10.25, (95% CI 10.14, 10.36, $P=10^{-16}$) well above the 2006 gradient's 95% confidence level. This is evidence that lifetime migration has diluted the latitudinal gradient from birthplace to the 2006 location. Fig. 4 also shows that the gradient is greater than the (overall) 2006 gradient for all ages up to 35, just after migration peaks.

3.4 Prevalence gradient stratified by migration status

Having demonstrated the latitudinal gradients, dominated by female relapsing remitting cases, are established at locations in the first 12 years of life, we now explore the relative contributions of different types of internal migrants to the latitudinal gradients.

Each NZ born case was classified at their location in 2006 by lifetime migration status, as being one of 4 categories:

1. Incident: Diagnosed in 2006 at location in 2006.
2. Migrated: Born in NZ in a region other than their 2006 resident region, possibly moved elsewhere then moved into the 2006 region.
3. Returned: Born in their 2006 residence region, left, and returned by 2006.
4. Remained: never resided elsewhere than their 2006 resident region.

The denominator is the total NZ born population in 2006. Fig. 5 shows the cumulative latitude gradients by migration type, revealing the complexity of the migration patterns of the NZ born sub cohort. It reveals that the gradient in 2006 is driven by those that 'remained' and those that 'returned' to each region (classes 3 and 4). This is most notable for region 6, Southland and Otago. The capital region 4, encompassing Wellington and Nelson, shows as expected, the greatest proportion of class 2 'migrated' cases, who have been attracted there either for employment or retirement.

4. Discussion

To summarise our findings, the latitude gradient for prevalence in the NZ born cohort is very similar to the wider cohort, but with a stronger signal at birthplace compared to the location in 2006. The gradient is driven by female relapsing remitting cases. The gradient is maximal at locations from birth

to 12 years old, and the 2006 gradient is driven by those who stayed in or returned to their region of birth. Hence the gradient (and risk exposure) is established early in life and potentially during gestation. The gradient is not materially affected by the process of migration and it is unlikely that the act of migration itself could have a causal effect on the multiple sclerosis gradient in NZ. Where migration does impact is in diluting the birthplace relationship to latitude in subsequent years.

This study is the first to demonstrate that the environmental drivers of multiple sclerosis risk, as manifested by the latitudinal gradient of multiple sclerosis prevalence are determined early in the life course. The multiple sclerosis latitudinal gradient has been demonstrated in multiple studies and has been shown to be increasing with time in both the Northern and Southern hemispheres (Simpson *et al.*, 2019). The multiple sclerosis latitudinal gradient is one of the strongest geo-epidemiological findings in multiple sclerosis risk studies with the risk of multiple sclerosis increasing significantly (more than 10 fold) the further the place of residence is from the equator (Simpson *et al.*, 2019). In our study the gradient of multiple sclerosis prevalence was present at birth in those born in NZ and that this gradient remained relatively stable until the age of about 12. This leads us to conclude that migration is acting to dilute the gradient after birth. This finding has significant implications as it indicates that the environment at or near the time of birth is a major factor determining the risk of developing multiple sclerosis and suggests that modifiable environmental factors for example sunlight exposure and vitamin D deficiency need to be addressed as early as possible in pregnancy and the neonatal period.

New Zealand's indigenous Māori have lower susceptibility to multiple sclerosis, consistent with other non-European ancestries. However, we and others have previously shown that genetic admixture and geographical distribution of Māori do not explain the latitudinal gradient of multiple sclerosis in NZ (Pearson *et al.*, 2014). Hence, in New Zealand, the observed latitudinal gradient is independent of the distribution of the indigenous population.

There are several caveats that need to be considered in this study. Firstly the latitude gradient for prevalence determined at birth required some assumptions to be made, particularly as the gradient could not be directly determined as the range of birth dates for participants in this study was from 1915 to 1990 and the population structure of New Zealand has changed significantly over this time period. We do not have data on how the whole NZ population has migrated throughout their lifetimes, as this longitudinal data-linkage is not available from historical censuses. In using the 2006 population for the denominators, we have been conservative in our overall earlier life course latitude gradient estimates, since we know that the NZ population grew from 2.2 million in 1956, the closest census to the mean year of birth, compared to the 2006 population of 4.2 million. The various regions of New

Zealand have grown at differing rates since 1956, notably the stronger growth in Auckland, in the north of the country. We thus know that our denominator did not change smoothly. We did a sensitivity analysis using the 1956 census data, and no material differences were seen in the latitudinal gradient analysis since the regions of the country with high prevalence (in the south) corresponded with the regions with least population change (data not shown). Thus, we can conclude that the shape of our latitude gradient is accurate, but we probably have underestimated the magnitude of the gradient.

Additionally, our analysis could not include all people born in New Zealand with multiple sclerosis as we received questionnaire responses from 73% of the 2917 cases that we identified, although we have previously shown that the 73% were representative of the whole NZ multiple sclerosis population (Taylor *et al.*, 2010a). Also, sensitivity analysis indicates that the NZ born group were different from the immigrating group on certain important factors and this has been noted before in other disease cohorts. However the NZ born group from the data available did not differ significantly from the whole population of people with multiple sclerosis in NZ in any meaningful way, indicating that the sub group analysed here is representative of the whole NZ multiple sclerosis population. Recording of lifetime residence relied on case's (or their carer's) recall which could have incurred recall bias. In the future, use of longitudinal research tools such as Stats NZ's Integrated Data Infrastructure (IDI) or record-linkage via New Zealand's National Health Index (NHI) could overcome recall bias, but unfortunately these tools do not currently allow retrospective identification of residential address prior to 2006. A further theoretical limitation is that we cannot consider emigration from NZ of people with MS after their diagnosis unless they returned and were resident in NZ for the 2006 prevalence study as we do not have any mechanism to identify them. However, it would seem unlikely that this would be many cases, or that they would differentially emigrate from any one region. Although our data strongly points to the gradient being present at birth it is possible, although unlikely, that fluctuations in the gradient may occur that are due to the epoch nature of the data collected and we simply do not have the power to further deconvolute this unlikely scenario.

These data support the observations suggested by month of birth studies that place of birth is an important determinant of subsequent multiple sclerosis risk (Dobson *et al.*, 2013) and by neonatal vitamin D studies (Munk Nielsen, 2017). However, there is no current evidence to suggest that maternal or neonatal UVR exposure has an independent effect from vitamin D deficiency alone. Other work has suggested that the risk of CIS is independently driven by both vitamin D and sunlight exposure (Lucas, 2011b), and that these factors could explain about 20% of the overall risk of developing a first clinical symptom suggestive of CNS demyelination (Lucas *et al.*, 2011b) and that

sunlight exposure throughout the life course is an important driver of multiple sclerosis risk although this was truncated to age 6. (Simpson *et al.*, 2019).

The finding that the latitude gradients by multiple sclerosis phenotype and sex at birth in this sub-cohort behaved in the same way, but with a stronger signal, than they did in the 2006 prevalence study (Taylor *et al.*, 2010a) indicates that the environmental drivers have a greater effect on females and those with relapse onset multiple sclerosis. The lack of gradient in progressive onset multiple sclerosis has also been reported from a large Australian CIS cohort (Taylor *et al.*, 2010b). The female predominance also is supported by the finding of an increased female to male ratio with latitude indicating strongly that females are more susceptible to latitude-based environmental factors (Trojano, 2012).

This study adds the major factor of time into the latitudinal gradient equation and suggests why studies at the time of diagnosis and at other prevalence points may not provide the answers as to what drives the gradient. As a consequence, our results focus the need for multiple sclerosis prevention strategies at significantly earlier time points, including potentially in utero. This does not affect the factors that are not necessarily latitudinally variant: such as genetic makeup, adolescent obesity, and smoking that have been independently associated with multiple sclerosis risk. The timing of EBV infection may have significant latitudinal variation with the majority of EBV infections in more tropical areas occurring early in life with less chance of developing infectious mononucleosis as compared to higher latitude countries (Lucas, 2011a).

The substantial difficulty in conducting research of this type with a disease of long latency such as multiple sclerosis where the mean age of onset in our cohort was 35 years and the mean age at prevalence was 53 means that a study that proves that multiple sclerosis can be prevented or reduced by interventions aimed at pregnant woman and neonates would require at least a 50 year observational study in a large population at risk. The development of biomarkers or identification of genetic markers that predict risk of multiple sclerosis later in life may allow identification of at-risk families or individuals at birth allowing a more focused but still long-term study to be undertaken.

Here, for the first time, we present epidemiological evidence that strongly suggests that the environmental drivers of the latitudinal gradient act at the earliest points in a person's life, perhaps even in utero or in the neonatal period. These data are of importance to those who are considering studies aimed at reducing multiple sclerosis risk by modification of environmental factors, which clearly need to be introduced at much earlier time points than have previously been considered.

Acknowledgements

We would like to thank Professor Ann Richardson for her contribution to the NZNMSPS.

Funding

This work was supported by a partnership grant from the NZ national MS society and the Health Research Council of New Zealand, Grant No HRC MS 05 524. Prof Sabel is supported by the Novo Nordisk Foundation Challenge Programme (grant NNF17OC0027864) for the Danish Big Data Centre for Environment and Health. Prof Taylor is supported by a Macquarie Foundation/Multiple Sclerosis Research Australia Senior Clinical Research Fellowship.

Competing interests

All authors declare that we have no significant competing financial, professional, or personal interests that might have influenced the performance or presentation of the work described in this manuscript

References

Alter M, Kahana E, Loewenson R. Migration and risk of multiple sclerosis. Neurology (Minn.) 1978; 28: 1089-1093.

Bentham G. Migration and Morbidity: Implications for Geographical Studies of Disease. Social Science & Medicine 1988; 26: 49-54.

Cabre P, Signate A, Olindo S, Merle H, Caparros-Lefebvre D, Béra O, Smadja D. Role of return migration in the emergence of multiple sclerosis in the French West Indies. Brain 2005; 128(Pt 12): 2899-2910.

Dobson R, Giovannoni G, Ramagopalan S. The month of birth effect in multiple sclerosis: systematic review, meta-analysis and effect of latitude. J Neurol Neurosurg Psychiatry 2013; 84(4): 427-432.

Ebers GC, Sadovnick AD. The geographic distribution of multiple sclerosis: a review, Neuroepidemiology 1993; 12 (1): 1-5.

Elian, M, Nightingale S, Dean G. Multiple sclerosis among United Kingdom-born children of immigrants from the Indian subcontinent, Africa and the West Indies. Journal of neurology, neurosurgery, and psychiatry 1990; 53(10): 906–911.

Flowerdew R, Manley D, Sabel C E. Neighbourhood Effects on Health: Does it Matter where you draw the Boundaries? Social Science & Medicine 2008, 66, 1241-1255.

Hägerstrand, T. What about people in regional science? Papers of the Regional Science Association 1970; 24 (1): 6–21

International Multiple Sclerosis Genetics Consortium. Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility. Science 2019; 365 (6460).

Jacquez G, Sabel CE, Shi C (2015) Genetic GIS: Towards a place-based synthesis of the genome, exposome and behavome, Annals Association of American Geographer 2015; 105 (3): 454-472.

Jia P, Lakerveld, J, Wu J, Stein A, Root, E, Sabel, C E, Vermeulen, R, Remais, J V, XI, C, Brownson R C, Amer S, Qian X, Wang L, Brownson R, Wu T, Xiao Q, Wang L, Sabel C E, Wang Y, Verschuren W M M, Wu T, Wang Y, James P. Top 10 research priorities in Spatial Lifecourse Epidemiology, Environmental Health Perspectives 2019; 127 (7).

Kurtzke, JF. Geography in Multiple Sclerosis, J Neurol 1977; 215 (1):1-26.

Kurtzke, JF. Multiple Sclerosis in Time and Space--Geographic Clues to Cause, J Neurovirol 2000; May 6; Suppl 2:S134-40.

Lucas RM, Ponsonby AL, Dear K, Valery P, Pender MP, Burrows JM, Burrows SR, Chapman C, Coulthard A, Dwyer DE, Dwyer T, Kilpatrick T, Lay ML, McMichael AJ, Taylor BV, van der Mei IA, Williams D. Current and past Epstein-Barr virus infection in risk of initial CNS demyelination, Neurology 2011a; 77(4):371-379.

Lucas RM, Ponsonby A-L, Dear K, Valery PC, Pender MP, Taylor BV, Kilpatrick TJ, Dwyer T, Coulthard A, Chapman C, van der Mei I, Williams D, McMichael AJ. Sun exposure and vitamin D are independent risk factors for CNS demyelination, Neurology 2011b; 76 (6): 540-548.

Mayer JD. Geographical Clues about Multiple Sclerosis, Annals of the Association of American Geographers 1981; 71(1): 28-39.

McLeod, JG, Hammond, SR, Kurtzke, JF. Migration and multiple sclerosis in immigrants to Australia from United Kingdom and Ireland: a reassessment. I. Risk of MS by age at immigration. Journal of Neurology 2011; 258(6): 1140–1149.

Munk Nielsen N, Munger KL, Koch-Henriksen N, Hougaard, D. M., Magyari, M., Jørgensen, K. T., Lundqvist, M., Simonsen, J., Jess, T., Cohen, A., Stenager, E., Ascherio, A. Neonatal vitamin

D status and risk of multiple sclerosis: A population-based case-control study. *Neurology*. 2017;88(1):44-51.

Munk Nielsen N, Corn G, Frisch M, Stenager E, Koch-Henriksen N, Wohlfahrt J, Magyari M, Melbye M. Multiple sclerosis among first- and second-generation immigrants in Denmark: a population-based cohort study. Brain 2019: 142(6): 1587-1597.

Pakpoor J, Schmierer K, Cuzick J, Giovannoni G, Dobson R. Estimated and projected burden of multiple sclerosis attributable to smoking and childhood and adolescent high body-mass index: a comparative risk assessment, Int J Epidemiol; 2020. Published online ahead of print.

Pearson JF, Alla S, Clarke G, Taylor BV, Miller DH, Richardson A, Mason DF. Multiple sclerosis in New Zealand Māori, Multiple Sclerosis Journal 2014; 20(14):1892-1895.

Ponsonby A-L, Lucas RM, Van Der Mei IA, Dear K, Valery PC, Pender MP, Taylor BV, Kilpatrick TJ, Coulthard A, Chapman C, Williams D, McMichael AJ, Dwyer T. Offspring number, pregnancy, and risk of a first clinical demyelinating event: the AusImmune Study. Neurology 2012; 78(12): 867-874.

Risco J, Maldonado H, Luna L, Osada J, Ruiz P, Juarez A and Vizcarra D. Latitudinal prevalence gradient of multiple sclerosis in Latin America, Multiple Sclerosis 2011; 17 (9): 1055-1059.

Rothman KJ. Causes. Am. J. Epidemiol 1976; 104: 587–592

Sabel CE, Gatrell AC, Löytönen M, Maasilta P, Jokelainen M. Modelling Exposure Opportunities: Estimating Relative Risk for Motor Neurone Disease in Finland, Social Science & Medicine 2000; 50: 1121-1137.

Sabel CE, Boyle PJ, Löytönen M, Gatrell AC, Jokelainen M, Flowerdew R Maasilta P. The spatial clustering of Amyotrophic Lateral Sclerosis in Finland at place of birth and place of death, American Journal of Epidemiology 2003; 157 (10): 898-905.

Simpson Jr S, Blizzard L, Otahal P, Van der Mei I, Taylor B. Latitude is significantly associated with the prevalence of multiple sclerosis: a meta-analysis, J Neurol Neurosurg Psychiatry 2011; 82(10):1132-41.

Simpson S Jr, van der Mei I, Lucas RM, Ponsonby A-L, Broadley S, Blizzardd L, Ausimmune/AusLong Investigators Group, Taylor B. Sun Exposure across the Life Course Significantly Modulates Early Multiple Sclerosis Clinical Course. Front Neurol. 2018; 9:16

Simpson Jr S, Wang W, Otahal P, Blizzard L, Van der Mei I, Taylor B. Latitude continues to be significantly associated with the prevalence of multiple sclerosis: an updated meta-analysis, J Neurol Neurosurg Psychiatry 2019; 90(11):1193-1200.

Taylor BV, Pearson JF, Clarke G, Mason D, Abernethy D, Willoughby E, Sabel C E (2010a) MS prevalence in New Zealand, an ethnically and latitudinally diverse country, Multiple Sclerosis 2010a; 16 (12):1422-31.

Taylor BV, Lucas RM, Dear K, Kilpatrick TJ, Pender MP, van der Mei IA, Chapman C, Coulthard A, Dwyer T, McMichael AJ, Valery PC, Williams D, Ponsonby AL. Latitudinal variation in incidence and type of first central nervous system demyelinating events, Multiple Sclerosis 2010b; 16(4):398-405.

Trojano M, Lucchese G, Graziano G, Taylor BV, Simpson S Jr, Lepore V, Grand'Maison F, Duquette P, Izquierdo G, Grammond P, Amato MP, Bergamaschi R, Giuliani G, Boz C, Hupperts R, Van Pesch V, Lechner-Scott J, Cristiano, E Fiol M, Oreja-Guevara C, Saladino M-L, Verheul F, Slee M, Paolicelli D, Tortorella C, D'Onghia M, Iaffaldano P, Direnzo V, Butzkueven H. Geographical variations in sex ratio trends over time in multiple sclerosis. PLOS One 2012; 7(10): e48078

van der Mei I, Lucas RM, Taylor BV, Valery PC, Dwyer T, Kilpatrick TJ, Pender MP, Williams D, Chapman C, Otahal P, Ponsonby AL. Population attributable fractions and joint effects of key risk factors for multiple sclerosis, Multiple Sclerosis 2016; 22(4):461-469

Wallin MT, Page WF, Kurtzke JF. Multiple sclerosis in US veterans of the Vietnam era and later military service: race, sex, and geography. Ann Neurol. 2004; 55(1):65-71.

Wallin MT, Page WF, Kurtzke JF. Migration and multiple sclerosis in Alaskan military veterans, J Neurol 2009; 256: 1413-1417

Wild, CP. Complementing the genome with an "exposome": The outstanding challenge of environmental exposure measurement in molecular epidemiology. Cancer Epidemiology Biomarkers & Prevention 2005;14 (8): 1847–50.

Figure legends

Figure 1. Aggregated census regions for New Zealand (NZ): all region population weighted centroids (PWCs) refer to the NZ statistical regions; aggregated PWCs refer to the combined regional PWCs as defined in the NZNMSPS.

Figure 2. Age adjusted latitude gradients for prevalence of multiple sclerosis for the Total cohort and for the NZ Born sub-cohort at prevalence day in 2006 and birthplace location.

Figure 3. Latitude gradient for prevalence of multiple sclerosis for the NZ born sub-cohort by sex and onset type.

Figure 4. Latitude gradient for prevalence of multiple sclerosis in NZ born sub-cohort with location at their birthday for ages up to 80 years. The horizontal line is the gradient for the NZ born sub-cohort in 2006, with 95% confidence intervals.

Figure 5. NZ born multiple sclerosis cases usually resident in NZ in 2006 by migration status Prevalence is with respect to NZ born Population in 2006. On the right: Cumulative latitude gradients with standard errors, * $P < 0.05$. *** $P < 0.001$.

STROBE statement: Reporting guidelines checklist for cohort, case-control and cross-sectional studies

SECTION	ITEM NUMBER	CHECKLIST ITEM	REPORTED ON PAGE NUMBER:
TITLE AND ABSTRACT			
	1a	Indicate the study's design with a commonly used term in the title or the abstract	2
	1b	Provide in the abstract an informative and balanced summary of what was done and what was found	2
INTRODUCTION			
Background and objectives	2	Explain the scientific background and rationale for the investigation being reported	3-5
	3	State specific objectives, including any pre-specified hypotheses	5
METHODS			
Study design	4	Present key elements of study design early in the paper	5-8
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5-8
Participants	6a	Cohort study—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up Case-control study—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls Cross-sectional study—Give the eligibility criteria, and the sources and methods of selection of participants	5-6
	6b	Cohort study—For matched studies, give matching criteria and number of exposed and unexposed Case-control study—For matched studies, give matching criteria and the number of controls per case Variables	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	5-8

SECTION	ITEM NUMBER	CHECKLIST ITEM	REPORTED ON PAGE NUMBER:
Data sources/measurements	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group.	5-6
Bias	9	Describe any efforts to address potential sources of bias.	5-6
Study size	10	Explain how the study size was arrived at	5-6
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why .	6-8
Statistical methods	12a	Describe all statistical methods, including those used to control for confounding	6-8
	12b	Describe any methods used to examine subgroups and interactions	6-8
	12c	Explain how missing data were addressed	5-8
	12d	Cohort study—If applicable, explain how loss to follow-up was addressed Case-control study—If applicable, explain how matching of cases and controls was addressed Cross-sectional study—If applicable, describe analytical methods taking account of sampling strategy	
	12e	Describe any sensitivity analyses	6-8
RESULTS			
Participants	13a	Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	7
	13b	Give reasons for non-participation at each stage	
	13c	Consider use of a flow diagram	
Descriptive Data	14a	Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	5-6
	14b	Indicate number of participants with missing data for each variable of interest	5-7
	14c	Cohort study—Summarise follow-up time (eg, average and total amount)	
Outcome Data	15*	Cohort study—Report numbers of outcome events or summary measures over time	7

SECTION	ITEM NUMBER	CHECKLIST ITEM	REPORTED ON PAGE NUMBER:
		Case-control study—Report numbers in each exposure category, or summary measures of exposure Cross-sectional study—Report numbers of outcome events or summary measures	
Main Results	16a	Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g. 95% confidence interval). Make clear which confounders were adjusted for and why they were included	8-11
	16b	Report category boundaries when continuous variables were categorized	8-11
	16c	If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	8-11
	16d	Report results of any adjustments for multiple comparisons	8-11
Other Analyses	17a	Report other analyses done—e.g. analyses of subgroups and interactions, and sensitivity analyses	8-11
	17b	If numerous genetic exposures (genetic variants) were examined, summarize results from all analyses undertaken	
	17c	If detailed results are available elsewhere, state how they can be accessed	
DISCUSSION			
Key Results	18	Summarise key results with reference to study objectives	11-12
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	12-13
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	11-13
Generalisability	21	Discuss the generalisability (external validity) of the study results Other information	13
FUNDING			
	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	13

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

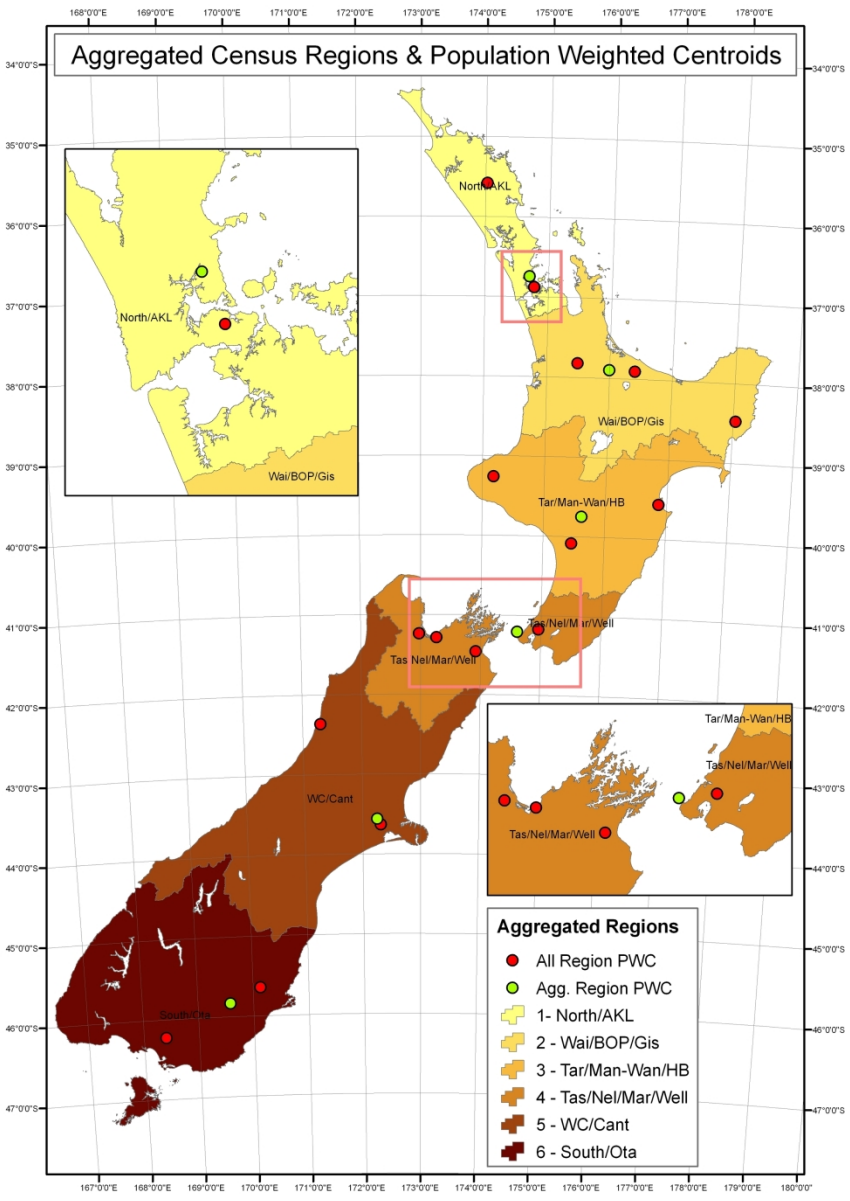


Figure 1. Aggregated census regions for New Zealand (NZ): all region population weighted centroids (PWCs) refer to the NZ statistical regions; aggregated PWCs refer to the combined regional PWCs as defined in the NZNMSPS.

209x297mm (300 x 300 DPI)

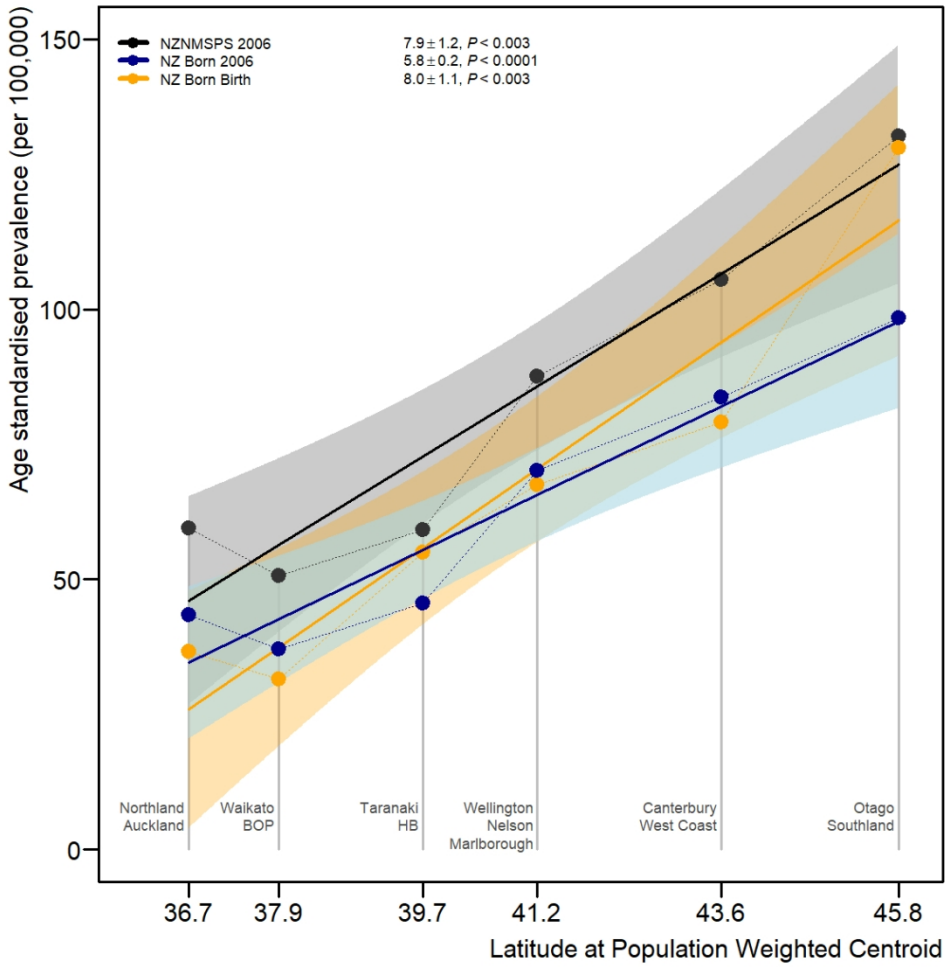


Figure 2. Age adjusted latitude gradients for prevalence of multiple sclerosis for the Total cohort and for the NZ Born sub-cohort at prevalence day in 2006 and birthplace location.

101x101mm (300 x 300 DPI)

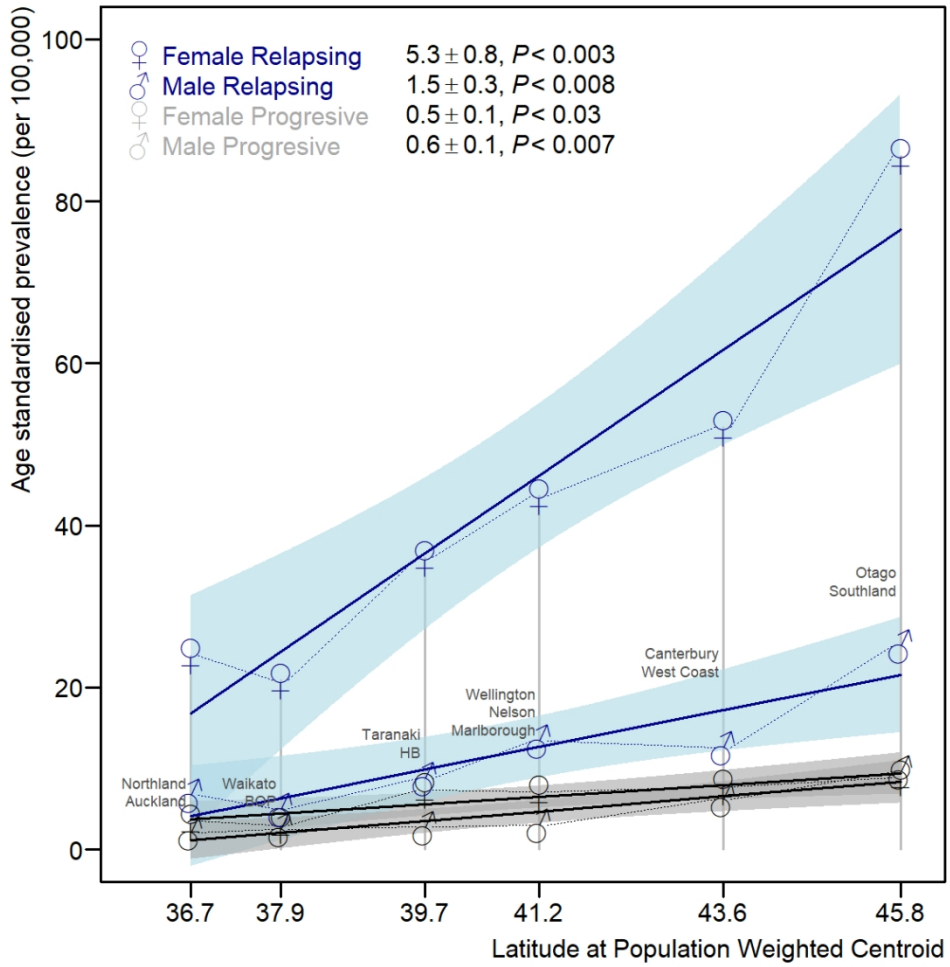


Figure 3. Latitude gradient for prevalence of multiple sclerosis for the NZ born sub-cohort by sex and onset type.

101x101mm (300 x 300 DPI)

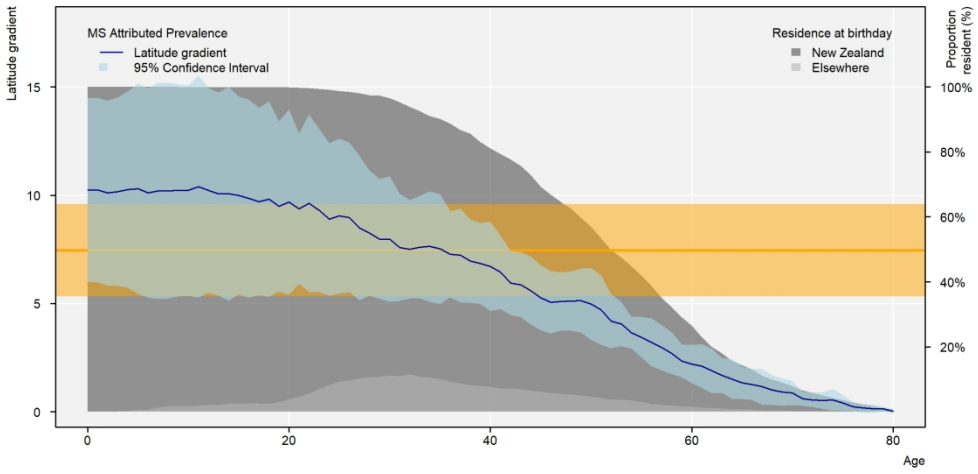


Figure 4. Latitude gradient for prevalence of multiple sclerosis in NZ born sub-cohort with location at their birthday for ages up to 80 years. The horizontal line is the gradient for the NZ born sub-cohort in 2006, with 95% confidence intervals.

203x101mm (300 x 300 DPI)

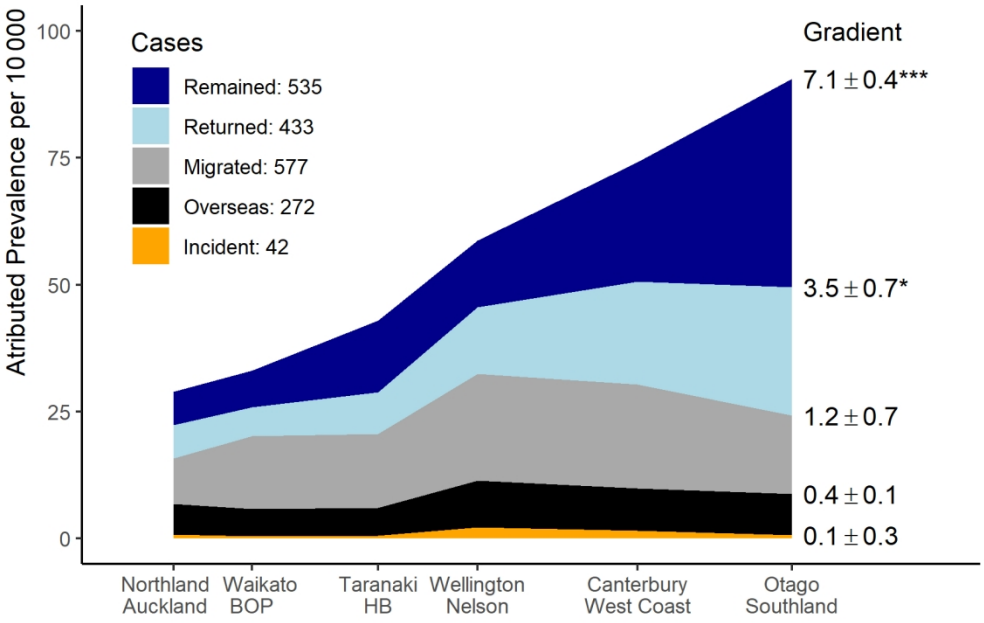


Figure 5. NZ born multiple sclerosis cases usually resident in NZ in 2006 by migration status
Prevalence is with respect to NZ born Population in 2006. On the right: Cumulative latitude gradients with standard errors, * P < 0.05. *** P < 0.001.

152x101mm (300 x 300 DPI)

	Born elsewhere	Born NZ	NZ birth to diagnosis	NZNMSPS
n	272	1587	1197	2917
Year of birth	*P* = 0.002	*P* = 0.163	*P* = 0.057	
Q1	1944	1945	1945	1946
Median	1952	1954	1953	1954
Q3	1960	1962	1962	1963
Sex	*P* = 0.942	*P* = 0.515	*P* = 0.662	
Female, *n* (%)	205 (75.4)	1206 (76.0)	907 (75.8)	2190 (75.1)
Male, *n* (%)	67 (24.6)	381 (24.0)	290 (24.2)	727 24.9)
Onset type				
Relapsing, *n* (%)	220 (80.9)	1315 (82.9)	981 (82.0)	2388 (81.9)
Progressive, *n* (%)	50 (18.4)	262 (16.5)	207 (17.3)	459 (15.7)
Other, *n* (%)	2 (0.7)	10 (0.6)	9 (0.8)	70 (2.4)
Location prevalence day	*P* = 0.002	*P* = 0.474	*P* = 0.143	
North 1, *n* (%)	87 (32.0)	334 (21.0)	228 (19.0)	819 (28.1)
2, *n* (%)	37 (13.6)	189 (11.9)	149 (12.4)	334 (11.5)
3, *n* (%)	26 (9.6)	178 (11.2)	146 (12.2)	279 (9.6)
4, *n* (%)	53 (19.5)	287 (18.1)	211 (17.6)	505 (17.3)
5, *n* (%)	46 (16.9)	364 (22.9)	278 (23.2)	598 (20.5)
South 6, *n* (%)	23 (8.5)	235 (14.8)	185 (15.5)	382 (13.1)

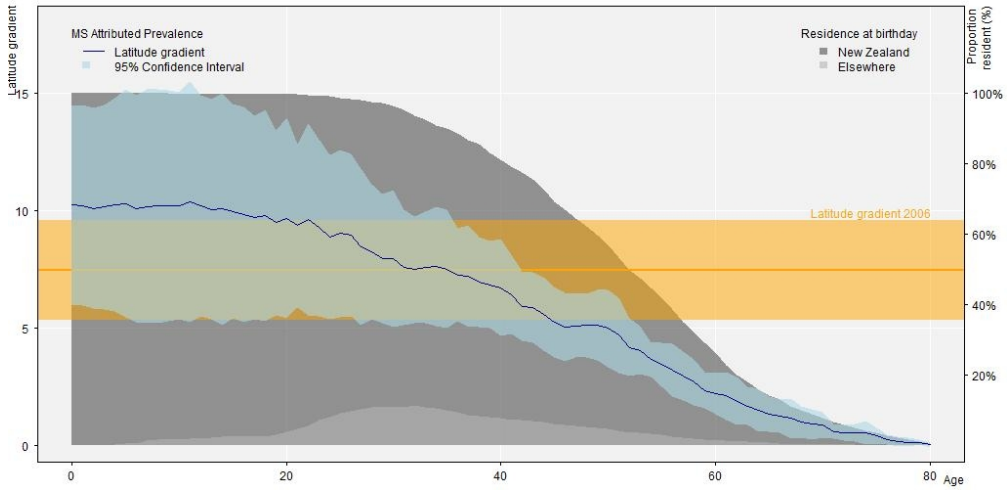
Table 1 Cases for the whole and subcohorts of NZNMSPS, by birth year, sex, onset type and geographic region

P-values calculated for subsample against entire NZNMSPS cohort, Wilcoxon tests for Year, Fisher's exact tests for sex and onset type (omitting 'other') and Kolmogorov-Smirnov tests on cumulative location.

Cohort	Raw prevalence			Age adjusted prevalence		
	Gradient	95% CI	*P*	Gradient	95% CI	*P*
All	9.50	5.93–13.07	0.0018	8.89	5.15–12.62	0.0027
NZ Born 2006	7.46	5.33–9.58	0.0006	6.96	4.23–9.70	0.0021
NZ Born Birth	10.25	6.00–14.50	0.0026	9.97	5.71–14.23	0.0029

Table 2 Raw and age adjusted latitude gradients for prevalence of multiple sclerosis

Latitude gradients of prevalence in each cohort with 95% confidence intervals with *P*-value testing against zero gradient, all gradients significant at the Bonferroni multiple testing level ($P = 0.017$).



361x180mm (72 x 72 DPI)

A robust latitudinal gradient exists in both northern and southern hemispheres for multiple sclerosis prevalence. Where a person lives is their strongest environmental risk factor for the disease. Sabel *et al*. show that risk is established at or near birth, with significant implications for the design of prevention studies.