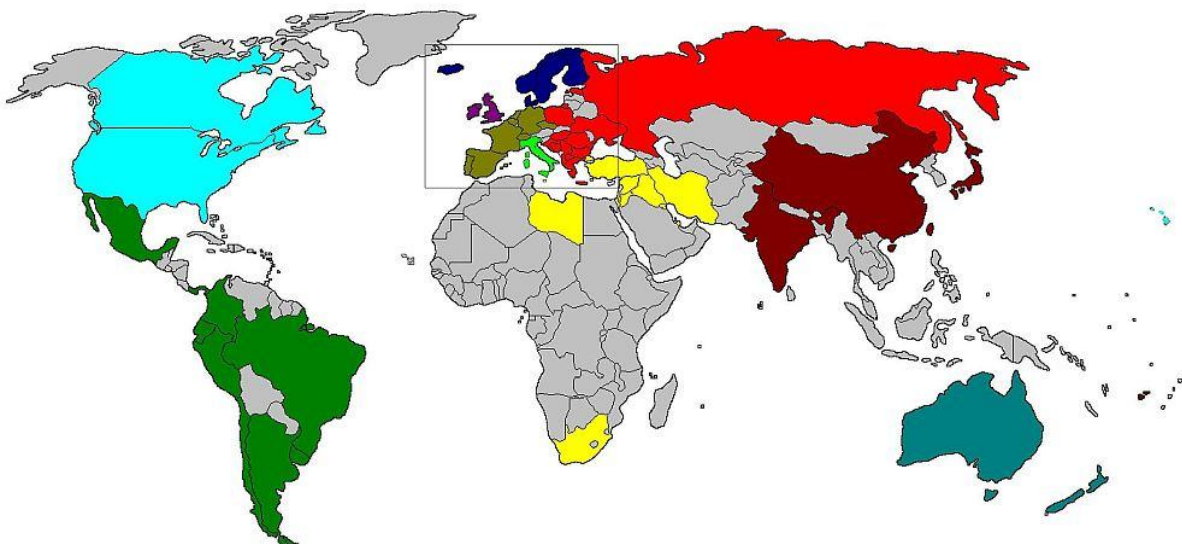


In establishing the regions used for the regional sub-analyses, some classifications were simple, based largely upon physical boundaries and geographic proximity, while others were more complex, grouped together on the basis of cultural/historical reasoning, or genetic homogeneity. Some of the classifications, particularly the separate groups for the UK and Scandinavia from the rest of Western Europe, are distinct from some studies<sup>1 2</sup> but are in harmony with others<sup>3-6</sup>, and are to our thinking appropriate divisions of populations which are particularly distinct, both genetically and culturally. Similarly our division of the Italian region from the rest of Western Europe is novel from some studies<sup>2-5</sup> but is similar to grouping done elsewhere<sup>1 6</sup> – here again, this division is appropriate, given the distinct epidemiology noted in the region<sup>7-15</sup>, but also because of the novel inverse gradient in this region found in our analysis.

The regions will be discussed in a different order from that used in the body of the text, saving the regions of Europe for last, given the interrelatedness of the decision-making for these. Regions at the global level are illustrated in Map 1. Regions within Europe are illustrated in Map 2.

**Map 1. World map showing the 10 study regions: aqua=Australasia; fuchsia=UK region; dark blue=Scandinavia & North Atlantic; green-brown=Atlantic & Central Europe; bright green=Italian region; red=Eastern Europe region; light blue=North America; green=Latin America & the Caribbean region; yellow=Middle East & Africa region; maroon=Asia & Pacific Islands region; grey=nations/regions without published prevalence data. Square boundary demarcates Europe region shown in Map 2.**



### **Australasia region (aqua)**

The Australasia region was demarcated as its own region, in keeping with virtually all reviews and meta-analyses which divided analyses into regions<sup>2-6</sup>. The logic behind this is principally a matter of geography and genetics – the location of the two nations of Australia and New Zealand naturally suggests their grouping, and their genetic and cultural distinctiveness from the surrounding region precludes their being grouped with others.

### **Asia & Pacific Islands region (maroon)**

This region was designated on the basis of the significant genetic and cultural differences between the nations of Asia and the Pacific Islands, here including Fiji, India, Japan, the People's Republic of China, and the Republic of China (Taiwan), and the rest of the world and this grouping is largely identical to that used elsewhere<sup>2-6</sup>.

### **North America region (light blue)**

The North America region was defined solely as the nations of Canada and the United States of America (USA), in contradistinction to previous works<sup>3,4</sup> which also included Mexico and the nations of the Caribbean, though in keeping with others<sup>2,6</sup>. We decided to evaluate the North America region as Canada and the USA alone given the pattern of founding and resultant genetic/cultural characteristics of the nations of the Americas. Canada and the USA have largely British foundations, as well as large components of French-descent, while Central and South America were almost entirely colonized by the Spanish and Portuguese. The south-western areas of the USA (Colorado, New Mexico, southern California, Texas) have significant influence from Latin America, not merely by virtue of their former inclusion in the Spanish Empire and subsequent independent nation of Mexico, but also the modern-day immigration from Latin America. However the predominantly-American culture of these areas and significant non-Hispanic populations therein, as well as a desire to maintain regional divisions at the national-level wherever possible left these to be included with the rest of the USA and thus, with the North America region.

### **Latin America and Caribbean region (green)**

This regional definition follows from the preceding North America region, encompassing the other nations of the Americas in Central and South America, and the Caribbean: Argentina, Brazil, Chile, Colombia, Ecuador, the French West Antilles of France, Mexico, Panama, Peru and Uruguay. As noted above, the nations of Central and South America are linked by their colonization by Spain and Portugal, giving them distinct genetic and cultural characteristics from the rest Canada and the USA, alongside significant homology with one another. The major exception to this Hispanic descent is in the French West Antilles, an overseas department of France. The significant non-European population, largely of African descent, renders this nation distinct from both North America and Latin America. However, on the basis of geography, and previous studies<sup>35,36</sup>, it was decided to include the French West Antilles in the grouping with Latin America.

### **Middle East & Africa region (yellow)**

Here again, the distinctiveness of the populations on the basis of genetic and cultural factors necessitate the separation of the nations in this region, Iran, Iraq, Israel, Jordan, Kuwait, Libya, Malta, Qatar, Saudi Arabia, South Africa, and Turkey, from their neighbours in Europe and Asia. While naturally the nations of the Middle East, Malta and Libya are quite distinct from South Africa, it was nonetheless necessary given the paucity of prevalence estimates in the latter. The two populations are not irreconcilable and certainly the populations of South Africa are not dissimilar from the nations of North Africa, and the latter are culturally linked to the nations of the Middle East.

This grouping is similar to that used elsewhere<sup>2-5</sup>, but is distinct in that we have included Turkey with the Middle East, rather than Europe. Turkey is a bridge between Europe and the Middle East, both physically and to varying extent, genetically and culturally. It was decided however that the ties binding Turkey to the Middle East were stronger, by virtue of the much greater cultural homology between these two<sup>16</sup> than between Turkey and the Eastern Europe region, with which Turkey would otherwise be grouped, as well as by racial/genetic groupings, again with much greater homology between Turkey and the Middle East, than with Eastern Europe.

A further significant change from the allocations found elsewhere is our inclusion of the nation of Malta with the Middle East & Africa region. While the Maltese population shows some homology with Sicily and Italy<sup>17</sup>, other work indicates a greater influence from the other side of the Mediterranean, in Tunisia, and directly from Phoenicia<sup>18</sup>. Certainly this hypothesis is borne out better by history, wherein Malta was, like Carthage, a major Phoenician settlement<sup>16 18</sup>. A further indication is that the inclusion of Malta in the analysis with the Italian region abrogates the strong inverse association observed there, suggesting a different relationship in Malta from Italy and Corsica.

## Europe

In evaluating Europe, we sought to take into account genetic and cultural differences to a much greater extent than geography, which necessarily becomes less prominent in the smaller area in question. An additional factor was the distribution of prevalence estimates – given the much greater number of studies done in Europe, finer divisions could be made to divide the populations more properly in consideration of their historical and genetic divisions, as well as to better capture unique associations which were present in the different regions. While we did evaluate the whole of Western Europe in keeping with work elsewhere<sup>2</sup>, we also evaluated the regions, demonstrating the presence of independent and distinct epidemiological associations which would otherwise be obscured in the aggregate due to an “averaging” of the individual effects.



the same time as Shetland, is much closer genetically to Scotland than Scandinavia than is Shetland<sup>22</sup>, possibly due to its much closer proximity to Scotland.

An argument could have been made for excluding Finland from the Scandinavian analysis. While certainly they are geographically and culturally related, the population of Finland are genetically quite distinct from the other Scandinavian nations<sup>23</sup> and the rest of Western Europe<sup>24</sup>, representing a complex genetic mixture due to their location between Western and Eastern Europe<sup>25</sup>. However, given the aforementioned geographic and cultural links, as well as the difficulty of then placing Finland, and particularly, given the population mixing found in the northern latitudes of Sweden, Norway and Finland, where all populations include some component of Finn<sup>26 27</sup>, we elected to leave Finland with the rest of the Scandinavia & North Atlantic region.

### **Italian region (bright green)**

This region was defined due to the need to evaluate Italy independently from the rest of Europe, as well as geographical considerations, and is in keeping with some previous reviews<sup>1 6</sup>. Italy has some genetic homology with Spain and Portugal<sup>28</sup> and the expansionism of the Roman Empire had lasting influences on the culture and language of much of Western Europe<sup>16</sup>. However the unique epidemiology of MS in the Italian region<sup>7-15</sup>, particularly Sardinia, necessitated its evaluation independently.

The inclusion of peninsular and insular Italy necessitated the additional inclusion of the island of Corsica. While Corsica is part of France, this only goes back to the mid 18th century<sup>16</sup> and indeed genetic analysis shows that Corsicans are more closely related to Sardinians than to the French<sup>29</sup>, which is to be expected given the history, as well as the physical proximity.

### **Eastern Europe (red)**

This region was demarcated on both genetic and historical-cultural bases, the latter being principally defined as those nations which were under Soviet influence over much of the 20<sup>th</sup> century<sup>16</sup> (excepting East Germany, which was evaluated with West Germany, and the Czech Republic, which was also evaluated with Germany) and is in keeping with work elsewhere<sup>2</sup>. This region thus included Albania, Bosnia-Herzegovina, Bulgaria, Croatia, Estonia, Greece, Hungary, Lithuania, Macedonia, Poland, Romania, Russia, Serbia, Ukraine, and the country formerly known as Yugoslavia. While the nations in the region are naturally quite distinct, both genetically and culturally, encompassing a huge swath of peoples and territory, many are linked genetically<sup>28</sup> and share linguo-cultural systems of similar derivation<sup>16</sup>.

### **Atlantic & Central Europe (orange)**

This grouping was difficult, given the paucity of age-standardisable prevalence estimates in the two major constituent regions comprised of those nations on the coast of the Atlantic Ocean and English Channel – Belgium, France, the Netherlands, Portugal and Spain - and the Germanic nations – the Czech Republic, Germany, and Switzerland. The former share a number of linguo-cultural links inherent in their shared histories and physical proximity<sup>16</sup>, as well as some genetic overlap<sup>28 30-32</sup>. The latter are thoroughly intertwined, both genetically<sup>28</sup>

<sup>30-32</sup> as well as in their language, culture and history. Unfortunately, as it was desirous to be able to evaluate associations after age-standardisation, these two regions had to be consolidated. The two are not uniformly dissimilar however – France shows some genetic homology with the Germanic region<sup>28</sup>, and continental Western Europe shares a number of cultural and historical interactions<sup>16</sup>.

1. Rosati G. Descriptive epidemiology of multiple sclerosis in Europe in the 1980s: a critical overview. *Ann Neurol* 1994;36 Suppl 2:S164-74.
2. Koch-Henriksen N, Sørensen PS. The changing demographic pattern of multiple sclerosis epidemiology. *The Lancet Neurology* 2010;9(5):520-32.
3. Kurtzke JF. A reassessment of the distribution of multiple sclerosis. Part one. *Acta Neurol Scand* 1975;51(2):110-36.
4. Kurtzke JF. A reassessment of the distribution of multiple sclerosis. Part two. *Acta Neurol Scand* 1975;51(2):137-57.
5. Kurtzke JF. General features on the prevalence of multiple sclerosis. *J. Indian Med. Prof* 1964;11:4896-901.
6. Rosati G. The prevalence of multiple sclerosis in the world: an update. *Neurol Sci* 2001;22(2):117-39.
7. Rosati G, Pinna L, Granieri E, Aiello I, De Bastiani P, Tola R, et al. The distribution of multiple sclerosis in Sardinia. *Riv Patol Nerv Ment* 1977;98(1):46-64.
8. Granieri E, Rosati G. Italy: a medium- or high-risk area for multiple sclerosis? An epidemiologic study in Barbagia, Sardinia, southern Italy. *Neurology* 1982;32(5):466-72.
9. Casetta I, Granieri E, Marchi D, Murgia SB, Tola MR, Ticca A, et al. An epidemiological study of multiple sclerosis in central Sardinia, Italy. *Acta Neurol Scand* 1998;98(6):391-4.
10. Montomoli C, Allemani C, Solinas G, Motta G, Bernardinelli L, Clemente S, et al. An ecologic study of geographical variation in multiple sclerosis risk in central Sardinia, Italy. *Neuroepidemiology* 2002;21(4):187-93.
11. Rosati G, Aiello I, Pirastru MI, Mannu L, Sanna G, Sau GF, et al. Epidemiology of multiple sclerosis in Northwestern Sardinia: further evidence for higher frequency in Sardinians compared to other Italians. *Neuroepidemiology* 1996;15(1):10-19.
12. Grimaldi LM, Salemi G, Grimaldi G, Rizzo A, Marziolo R, Lo Presti C, et al. High incidence and increasing prevalence of MS in Enna (Sicily), southern Italy. *Neurology* 2001;57(10):1891-3.
13. Grimaldi LM, Palmeri B, Salemi G, Giglia G, D'Amelio M, Grimaldi R, et al. High prevalence and fast rising incidence of multiple sclerosis in Caltanissetta, Sicily, southern Italy. *Neuroepidemiology* 2007;28(1):28-32.
14. Massa R, Mamusa E, Sardu C, Ferrigno P, Aste R, Frau J, et al. Epidemiology of multiple sclerosis in south western Sardinia: nature versus nurture. *Multiple sclerosis* 2007;13(Supplement 1):S103.
15. Iuliano G, Napoletano R. Prevalence and incidence of multiple sclerosis in Salerno (southern Italy) and its province. *Eur J Neurol* 2008;15(1):73-6.
16. Davies N. *Europe: a history*. New York: Oxford University Press, Inc., 1996.
17. Capelli C, Redhead N, Romano V, Cali F, Lefranc G, Delague V, et al. Population structure in the Mediterranean basin: a Y chromosome perspective. *Annals of human genetics* 2006;70(Pt 2):207-25.
18. Zalloua PA, Platt DE, El Sibai M, Khalife J, Makhoul N, Haber M, et al. Identifying genetic traces of historical expansions: Phoenician footprints in the Mediterranean. *American journal of human genetics* 2008;83(5):633-42.
19. Roberts DF. Genetic affinities of the Shetland islanders. *Ann Hum Biol* 1990;17(2):121-32.



20. Helgason A, Hickey E, Goodacre S, Bosnes V, Stefansson K, Ward R, et al. mtDna and the islands of the North Atlantic: estimating the proportions of Norse and Gaelic ancestry. *American journal of human genetics* 2001;68(3):723-37.
21. Berry RJ, Muir VML. The natural history of man in Shetland. *Journal of Biosocial Science* 1975;7(3):319-44.
22. Goodacre S, Helgason A, Nicholson J, Southam L, Ferguson L, Hickey E, et al. Genetic evidence for a family-based Scandinavian settlement of Shetland and Orkney during the Viking periods. *Heredity* 2005;95(2):129-35.
23. McEvoy BP, Montgomery GW, McRae AF, Ripatti S, Perola M, Spector TD, et al. Geographical structure and differential natural selection among North European populations. *Genome Res* 2009;19(5):804-14.
24. Salmela E, Lappalainen T, Fransson I, Andersen PM, Dahlman-Wright K, Fiebig A, et al. Genome-wide analysis of single nucleotide polymorphisms uncovers population structure in Northern Europe. *PLoS ONE* 2008;3(10):e3519.
25. Lappalainen T, Hannelius U, Salmela E, von Döbeln U, Lindgren CM, Huoponen K, et al. Population structure in contemporary Sweden--a Y-chromosomal and mitochondrial DNA analysis. *Annals of human genetics* 2009;73(1):61-73.
26. Einarsdóttir E, Egerbladh I, Beckman L, Holmberg D, Escher SA. The genetic population structure of northern Sweden and its implications for mapping genetic diseases. *Hereditas* 2007;144(5):171-80.
27. Karlsson AO, Wallerstrom T, Gotherstrom A, Holmlund G. Y-chromosome diversity in Sweden - a long-time perspective. *European journal of human genetics : EJHG* 2006;14(8):963-70.
28. Cavalli-Sforza L. *The history and geography of human genes*. Princeton: Princeton University Press, 1994.
29. Grimaldi MC, Crouau-Roy B, Amoros JP, Cambon-Thomsen A, Carcassi C, Orru S, et al. West Mediterranean islands (Corsica, Balearic islands, Sardinia) and the Basque population: contribution of HLA class I molecular markers to their evolutionary history. *Tissue Antigens* 2001;58(5):281-92.
30. Novembre J, Johnson T, Bryc K, Kutalik Z, Boyko AR, Auton A, et al. Genes mirror geography within Europe. *Nature* 2008;456(7218):98-101.
31. Seldin MF, Shigeta R, Villoslada P, Selmi C, Tuomilehto J, Silva G, et al. European population substructure: clustering of northern and southern populations. *PLoS genetics* 2006;2(9):e143.
32. Heath SC, Gut IG, Brennan P, McKay JD, Bencko V, Fabianova E, et al. Investigation of the fine structure of European populations with applications to disease association studies. *European journal of human genetics : EJHG* 2008;16(12):1413-29.