REPORT

Relationships of geochemistry and multiple sclerosis

Karin Eliaeson, IVL
Marcus Liljeberg, IVL
Anders Björk, IVL
Ingrid Kockum, KI
Alasdair Skelton, SU
Eva Jirner Lindström, SGU

B1824
December 2008

The report approved:
2008-12-31

Lars-Gunnar Lindfors
Scientific Director

IVL Swedish Environmental Research Institute

Organization	Report Summary
IVL Swedish Environmental Research Institute Ltd.	**Project title**
Address P.O. Box 21060 SE-100 31 Stockholm	**Project sponsor** **IVL and SGU**
Telephone +46 (0)8-598 563 00	

Author

Karin Eliaeson, Marcus Liljeberg, Anders Björk, Ingrid Kockum, Eva Jirner Lindström, Alasdair Skelton

Title and subtitle of the report

Relationships of geochemistry and multiple sclerosis

Summary

Se below

Keyword

Medical geology, multiple sclerosis, zinc, natural background levels, multivariate analysis, epidemiology

Bibliographic data

IVL Report B1824

The report can be ordered via

Homepage: www.ivl.se, e-mail: publicationservice@ivl.se, fax+46 (0)8-598 563 90, or via IVL, P.O. Box 21060, SE-100 31 Stockholm Sweden

Index

Foreword

This work is financed by Swedish Geological Survey (SGU) under the Research and Development program 2007 (FoU) and by the "7% development fund" of IVL Swedish Environmental Research Institute Ltd. The Geological department of Stockholm University and Department of Clinical Neurosciences at Karolinska Institutet have contributed with valuable expertise through the participation of Professor Alasdair Skelton and Associate professor Ingrid Kockum.

Summary

The main aim of this study has been to investigate how registers and databases of geochemistry can be combined with registers of patient data in epidemiological studies. By testing the hypothesis that Multiple Sclerosis (MS) varies with geography and investigating if the variation can be explained by natural variability of zinc in different media, difficulties have been identified and recommendations for future epidemiological studies with similar scope are given.

Multiple sclerosis (MS) is a chronic neurological illness that affects nerve cells in the central nervous system (CNS). It belongs to a group of illnesses called autoimmune diseases where the immune system attacks the body's own tissue. The onset of autoimmune reactions is not fully understood. Autoimmune diseases are believed to be multifactorial where both intrinsic factors (e.g. genetics, age, hormones) and environmental factors (e.g. infections, diet, drugs, chemicals etc.) may contribute to the induction, development and progression of the disease. There is a general believe that the epidemiological pattern of MS vary with geography, but even though the systematic study of MS started in 1929 the comparison of prevalence studies over the world still is very difficult and the results are not reliable. Iron, zinc, manganese, copper and molybdenum are examples of important building blocks for almost all living organisms and are thus termed essential elements. They originally derive from the Earth's crust and are taken up in organisms from soil, air and water. For some metals, no biological, nutritional or biochemical function has been established (yet) and they are thus termed non-essentials. The level of exposure to essential and non-essential elements is of crucial importance for the effect on living organisms. A too high dose can be toxic while a too low dose of essential elements will cause deficiency and consequent higher vulnerability for the exposure to toxic compounds or non-essential elements. In this study we have initially focused to check if the occurrence of MS could be correlated to background levels of zinc (Zn) since zinc is an element that participates in several important reactions in the body.

We used the Swedish MS-register, which includes almost all MS-patients in Sweden. The best resolution on where the patients live is given on post code areas. Spatially distributed census data over postcode areas, valid for December 2005 and compiled by Official Statistics of Sweden, Statistics Sweden (SCB), were used in this study. Geochemical data from soil (till), stream-water and groundwater from the Swedish Geological Survey have been compiled into postcode areas.
The analyzed data on the distribution of MS-patients indicate that a geographical pattern could be found with higher prevalence of MS in the county of Västerbotten and clusters around larges cities. No north –south or east-west gradient of the prevalence was found. However, visual interpretation of prevalence measures is strongly biased towards large post code areas, masking the variation of prevalence measures of small areas. This effect is striking in larger cities always having a large number of small post code areas.

Combination of the patient registers and the geochemical registers was evaluated with multivariate analysis (MVA) and as a univariate study for zinc solely, but no correlation between the prevalence of MS and the occurrence of natural background levels of elements were found. Registers were analyzed both separately and together, but none of these models increased the degree to which variance was explained. This does not mean that no relationship between MS and geochemistry is possible but that correlations could not be found with the data, methods and models used in this project.

The most important conclusion from this study is that to combine patient data with any kind of exposure data with a geographical variation, the administrative division (i.e. parishes, post code areas etc) are less appropriate. Divisions with respect to natural (geographical) borders such as

catchment areas would be more useful for epidemiological purposes where a geographic component is of interest. To fulfil this, population data for catchment areas is needed. The density of the patient data and the exposure data is also of crucial importance. Moreover, there must be a variation in the exposure data large enough to result in a difference between areas. It is recommended that also the areas where no disease is found to be included in epidemiological studies. In these regions high or low levels of elements can also be present. The use of average values over districts is problematic. A high density of sampling in an area does not necessarily mean that the calculated mean value is representative for the whole area. How well an average value for a district describes the actual value depends both on the natural variability of substances in the media as well as the sampling density (i.e. high variance but many samples could give the same average value as low variance and few samples).

1 Introduction

The main aim of this study has been to investigate how registers and databases of geochemistry can be combined with registers of patient data in epidemiological studies. Many epidemiological studies conclude on results that might be just a consequence of gaps in data causing false patterns. Thus, the method for combining different databases is a crucial tool and very important for receiving sound results that are statistically significant. This study illuminates some of the problems that might arise in epidemiological studies with a geographic component. By testing the hypothesis that Multiple Sclerosis (MS) varies with geography and investigating if the variation can be explained by natural variability of zinc in different media, difficulties have been identified and recommendations for future epidemiological studies with similar scope are given.

Also, this study is an attempt to link traditional medical research and natural science in the interdisciplinary research field of "medical geology". Environmental and medical researchers would gain from cross disciplinary research to receive a more holistic perspective on effects on human health and environment from both naturally occurring and antrophogenically derived substances.

2 Background

2.1 Natural background levels and human exposure

Humans are exposed to natural background levels of elements by drinking water, bathing water, cooking water, food (crops), air (dermal uptake, inhalation of particles) and so on. Today it is difficult to investigate the exposure from food since we consume food from all over the world and only a small proportion comes from local crops and animals. The drinking water on the other hand, is at many sites over the world taken from local water reserves and the exposure of natural background levels in groundwater are more easily studied. In Sweden it is very common that people get their drinking water from private water wells in soil or rock. Approximately 1.2 million of the Swedish population (9 million) use drinking water from private wells (SSI, 2008). The municipal water is either taken from surface water or groundwater. Both types (private wells and municipal wells) will be steered by natural prevailing conditions but groundwater is more easily correlated to geology than surface water since it is less rapidly affected by anthropogenic sources of pollution.

Metals and other elements are concentrated in bedrock by various magmatic and metamorphic processes. It is widely recognized that many of the chalcophile (associated with sulfides) and siderophile (occur as the native metal and occur in the core) elements (such as Ni, Co, Pt, Pd and Au) are more likely to be associated with mafic rock types whereas concentration of many lithophile (associated with silicates and concentrated in the crust) elements (such as Li, Sn, Zr, U and W) are typically found in association with felsic or alkaline rock types. Some average concentration of elements in the three main groups of magmas is given in table 1. However, remobilization of elements by hydrothermal and sedimentary processes can concentrate elements to significant degree in both metamorphic and sedimentary rock types. This is common in, for example, ore-deposits.

Table 1. Average values of elements for three magma types. The clarke-value is a term that refers to the avarege crustal abundance. Table after Robb 2008.

Element ppm	Magma type Basalt	Andesite	Rhyolite	Clarke value
Li	10	12	50	20
Be	0,7	1,5	4,1	2,8
F	380	210	480	625
P	3200	2800	1200	1050
V	266	148	72	135
Cr	307	55	4	10
Co	48	24	4,4	25
Ni	134	18	6	75
Cu	65	60	6	55
Zn	94	87	38	70
Zr	87	205	136	165
Mo	0,9 - 2,7	0,8 - 1,2	1	1,5
Sn	0,9	1,5	3,6	2
Nb	5	4 - 11	28	20
Sb	0,1 - 1,4	0,2	0,1 -0,6	0,2
Ta	0,9		2,3	2
W	1,2	1,1	2,4	1,5
Pb	6,4	5	21	13
Bi	0,02	0,12	0,12	0,17
U	0,1 - 0,6	0,8	5	2,7
Th	0,2	1,9	26	7,2
Ag (ppb)	100	80	37	70
Au (ppb)	3,6		1,5	4
Pt (ppb)	17 - 30		3 - 12	10
S	782	423	284	260
Ge	1,1	1,2	1,0 - 1,3	1,5
As	0,8	1,8	3,5	1,8
Cd	0,02	0,02	0,2 - 0,5	0,2

The concentration of elements in groundwater can reflect the surrounding geology (bedrock and soil-type) but it is also steered by rates of weathering and infiltration of rainwater (groundwater renewal), rainwater chemistry, the amount of humic acids, clay minerals and other potential chelating agents. Groundwater chemistry is further dependant on frequency of withdrawal of water from wells. The chemistry of groundwater is complex since there are many controlling factors but it is likely to be similar for regions with the same prevailing geologic and environmental conditions. Guideline values for drinking water regarding health aspects are given out by, for example, WHO 2006B (international) and SOSFS 2003:17 (national, i.e. Sweden). However, even though we have information about the concentration of different elements in groundwater, little is known about the actual uptake of these elements in the human body. Even less is known about the single and synergetic effects (both positive and negative) on human health of overall groundwater chemistry. Guideline values are mostly given for maximum tolerable levels but very seldom for minimum levels. Moreover, we seldom analyze all elements that are present in the water which restricts our knowledge of overall groundwater chemistry. The selection of parameters in regular groundwater monitoring programmes differs between studies, for example, because analytical techniques for specific elements or complexes are unavailable.

Natural background levels of different medias are commonly monitored on a national basis. Reasons for monitoring differ from mineral exploration to environmental purposes. Lately, a lot of effort has been put into combination of national registers into international databases to cover larger areas. For example, the Geochemical atlas of Europe was published in 2005 (Salminen et al 2005; De Vos et al, 2006) covering geochemical data of topsoil, subsoil, humus, stream sediment, stream water and floodplain sediment in 26 countries. The combination of different databases is demanding since it requires that sampling and analyzing techniques are consistent so that results can be comparable. A lot of work still remains to be done before good comparable data between nations exists. Similarly, registers of patient data are set up nationwide. Also here we find many different registers with approximately the same purposes. The purpose of the register and the criteria for diagnosis might differ between registers and consequently also between nations. In fact it is rare to have nationwide registries of the prevalence of diseases, Scandinavia being an exception due to the ability to trace individuals due to unique identification numbers (Rosati 2001).

2.2 Elements essential for organisms

Iron, zinc, manganese, copper and molybdenum are examples of important building blocks for almost all living organisms and are thus termed essential elements. They originally derive from the Earth's crust and are taken up in organisms from soil, air and water. An element is considered essential when it is present in living matter, it is able to interact with living systems and a deficiency results in a reduction of a biological function that is preventable or reversible by physiological amounts of the element (ICMM, 2007). The essentiality of elements differs between species, for example, Ni, Cr and Se have been established essential for a limited number of species (ICMM, 2007). For some metals, no biological, nutritional or biochemical function has been established (yet) and they are thus termed non-essentials. The level of exposure to essential and non-essential elements is of crucial importance for the effect on living organisms. A too high dose can be toxic while a too low dose of essential elements will cause deficiency and consequent higher vulnerability for the exposure to toxic compounds or non-essential elements. Organisms can, to a certain degree, cope with a range of varying concentration of essential elements. Within this range, the organism can regulate and maintain the optimal level of essential elements even though the external circumstances vary. This range is called the Optimal Concentration Range for Essential Elements (OCEE) and levels below or above this range will cause deficiency or toxicity to the organism. Thus, many elements can be toxic to organisms if the level is too high even if the element is essential. Contrary, too low levels will cause a deficiency and could lead to substitution of the needed element by others with similar properties that fit into the cell structure. The level of tolerance/deficiency can vary between organisms and it can also be influenced by synergetic effects between other elements and chemical reactions in the body.

2.3 General background on Multiple Sclerosis

Multiple sclerosis (MS) is a chronic neurological illness that affects nerve cells in the central nervous system (CNS). It belongs to a group of illnesses called autoimmune diseases where the immune system attacks the body's own tissue. Autoimmunity is characterized by the reaction of cells (e.g. autoreactive T-lymphocytes) or products (e.g. autoantibodies) of the immune system against the organism's own antigens (autoantigens) which will result in clinical abnormalities (autoimmune diseases) (WHO, 2006A). There are to date more than 60 diseases that are suspected to be autoimmune. They are characterized by the inappropriate or excessive immune response against autoantigens leading to chronic inflammation, tissue destruction and/or dysfunction (WHO, 2006A).

In MS, the nerve tissue that surrounds the nerve cells is attacked. The tissue is called myelin and is necessary for fast transportation of the impulses from the nerves. The regeneration of this type of cell is very slow and for this reason the destruction of myelin is very serious. It will result in a cease or a reduction of the speed of the nerve impulses which finally will lead to the development of the clinical signs of MS (problems with vision, sensation, strength and coordination). The word multiple sclerosis means "many scars" and refers to all the scars that the illness causes in the central nervous system.

The onset of autoimmune reactions is not fully understood. Autoimmune diseases are believed to be multifactorial where both intrinsic factors (e.g. genetics, age, hormones) and environmental factors (e.g. infections, diet, drugs, chemicals etc.) may contribute to the induction, development and progression of the disease. The prevalence[1] of MS is approximately 60 cases per 100 000 and the incidence[2] is 3 cases per 100 000 per year (WHO 2006A) but these statistics vary considerable between countries. MS is more common among women than men and seems to be very uncommon among black people. MS is not hereditary but an increased risk of developing the disease can be inherited. Approximately 20 % of people with MS are closely related to another person suffering from MS. Many of the risk genes are the same for different autoimmune diseases, for example the HLA DRB1-gene on chromosome 6 is a risk gene for virtually all autoimmune diseases including both diabetes type-1 and MS (Olerup et al 1991, Kockum et al 1995). Recently additional shared susceptibility genes for type-1 diabetes and MS have been identified including IL7R, IL2R, CLEC16A, CD226, SH2B3 and MHC2TA (Swanberg, Lidman et al. 2005; Ahmed, Gyllenberg et al. 2006; Lundmark, Duvefelt et al. 2007; Todd, Walker et al. 2007; International Multiple Sclerosis Genetics Consortium (IMSGC) In press; International Multiple Sclerosis Genetics Consortium (IMSGC) In Press)

There is a general believe that the epidemiological pattern of MS vary with geography, but even though the systematic study of MS started in 1929 the comparison of prevalence studies over the world still is very difficult and the results are not reliable (Rosati, 2001). The variability in surveyed population size, age structures, ethnic origin and the complexness of the actual onset of MS obstruct efforts to map its prevalance. MS most commonly strikes young adults between 20 and 40 years of age and the progression of the disease can be divided into three diagnoses: "relapsing remitting" (RRMS), "primary progressive" (PPMS) and "secondary progressive" (SPMS) (WHO, 2006A). In fact, some believe that MS is actually more than one disease. However because the frequency of risk versions of the susceptibility genes, such as HLA DRB1, are very similar in the different forms of MS they may just be different manifestations of the same disease. The unspecific onset date of the illness makes it difficult to find causative factors. Many hypotheses have arisen over time for example; degree of solar radiation (Ponsonby et al., 2002), intake of vitamin D (Van Amerongen et al., 2004, Ponsonby et al., 2005, Munger et al., 2004), dietary intake of fat (Zhang et al, 2000), tobacco smoke (Riise et al., 2003, Nielsen et al, 2007), Epstein Barr virus (Nielsen et al, 2007), exposure to organic solvents (Riise et al, 2002) and so on. As will be mentioned below, some researchers also speculate that exposure to metals can influence the evolution of MS.

As already mentioned, MS is thought to vary geographically. Theories of a general north-south gradient with an increase in incidence with the distance from the equator have been presented (Kurtzke, 1980). On the contrary, studies from Norway show a south – north decreasing gradient. This has been explained by the high proportion of Samis in the north of Norway that seems to be a population resistant to MS (Rosati, 2001). Also, theories of higher prevalence in inland areas in Norway as a consequence of low fish diet, have been discussed (Swank et al, 1952). Some researchers believe that MS is epidemic and that the central source is from the southern inland lake region of Sweden (Kurtzke, 2006). Migration from a low risk area to a high risk area in young age has been proposed to result in higher risk of developing the disease (Compston 1997; Marrie 2004).

[1] The prevalence is the number of people who have the disease and so is determined by the incidence and duration of the illness (WHO 2006A).

[2] The incidence of a disease is the number of newly diagnosed individuals that occur in a given time period (WHO 2006A).

Landtblom et al (2005) presented a study of the distribution of MS in Sweden. This study concluded that the maximum rates for development of MS in the early prevalence were found around the two major lakes in south central Sweden as well as for one region on the northern shore of the Bay of Bottnia, and another off the Bay north of Stockholm. They conclude that the distribution MS in Sweden could not only be due to genetics.

2.4 The significance of zinc and other metals in MS

In this study we have initially focused on zinc (Zn) since zinc is an element that is involved in important reactions in the body during the course of the disease. Furthermore, another Swedish study has shown that the onset of childhood diabetes can be influenced by the amount of zinc in groundwater (Haglund et al, 1996). This is interesting since MS and childhood diabetes have been identified to share some of the same risk genes. Thus, it is worth testing if MS also could be correlated to amount of zinc in groundwater. The databases used in this study contained information about several other elements than zinc. For this reason, multivariate statistical methods were used in this study to investigate not only the possible influence from zinc but also from the other substances. The importance of some essential elements in the human body (that occur naturally in groundwater) is discussed below.

Zinc participates in several important reactions in the body. For example, the functionality of more than 70 enzymes is dependent on the presence of zinc. The immune defence system is dependent on zinc and it is also a constituent in the respiration of the enzymes. It facilitates the growth of nails, hair, bone and participates in the healing of wounds. Zinc is also part of the enzyme amylase that transforms starch and is active when proteins and nucleine acid are synthesised. Furthermore, zinc is a prerequisite for the uptake of iron and copper and the construction of blood, it is necessary for normal function of the prostate and for optimal use of vitamin A. Zinc is common in liver, lamb, meat, egg, milk, carrots, nuts, rice and others. Recommended daily intake of zinc varies between different national recommendations but values are approximately: Children 2-7 mg, women 7-9 mg (pregnant and breast-feeding 9/11 mg), men 9-12 mg (WHO, 2001).

Zinc has the same electronic configuration and similar radius as cadmium (Cd) and mercury (Hg). This means that if zinc levels are too low in the body, cadmium and mercury can substitute for zinc in the cell structure. Zinc is essential for life while cadmium and mercury are classified as non-essentials. Lead (Pb), another non-essential, is also a possible substitute for Zn in the cell structure. Zinc deficiency could consequently result in poisoning from the substituting elements and reduced function of the many important zinc-dependent enzymes. For MS, an interesting reaction is the transport of carbon dioxide in the blood vessels that surround the CNS that is facilitated by the zinc-dependent enzyme carbonic anhydrase. The carbonic anhydrase dehydrates and compacts myelin by pumping ions and water into the cell. Magnesium is needed to pump calcium (Ca) and other ions out of membranes and cells. Thus, if magnesium deficiency occurs it will result in calcium overload and imbalances of water and ions and if Zn deficiency occurs it will inhibit the metabolism/ natural function of zinc dependent enzymes. Both of these processes are important in MS. The action of Matrix metallo proteinases (MMPs) are also thought to be important in the development of MS. These proteins are a family of zinc- and calcium dependent secreted or membrane bound endopetidases, Oathological expression and activation have been associated with MS (Wagner, Breyholz et al. 2006). An imbalance in MMP proteins or their activation can lead to accelerated destruction of connective tissues which is associated with the pathology of MS (Borkakoti 2004). MMPs and related enzymes can also process a number of cell surface cytokine receptors and other soluble proteins. For example they are involved in the release of TNFa, a pro-inflammatirt cytokine, from its membrane bound precursor, TNF a is thought to be important in the activation of the immune system and affects the onset of MS (Chandler, Miller et al. 1997). T-helper cells in the immune system are sometimes classified into Th1 and Th2 cells. Autoimmunity can arise though an imbalance between the activity of Th1 and Th2 cells. Zinc is one of the nutrients that has been shown to affect the Th1/Th2 balance (Kidd 2003).

Some other metals are important for the function of transport mechanisms and chemical regulation in the organism. For example iron is an important element for the transportation of oxygen from the lungs to the muscles. The most important group of iron-binding proteins contains the heme molecules, all of which contain iron at their centers. People with iron deficiency cannot produce an adequate amount of hemoglobin to meet their body's oxygen-transport needs. However, if iron overload becomes severe the condition is diagnosed as hemochromatosis which can result in serious damage to the body's tissues, including cirrhosis of the liver, heart failure, diabetes, abdominal pain, and arthritis. A recessive genetic mutation can put some people (e.g., those of Irish or Celtic descent) at a higher risk for developing hemochromatosis (Casiday and Frey, www.chemistry.wustl.edu).

Copper is another divalent cat ion with important biological functions. The essentiality of copper arises from its specific incorporation into a large number of enzymatic and structural proteins. The role of copper in oxidation/reduction enzyme activities is a consequence of its ability to function as an electron transfer intermediate. Thus copper is present in enzymes involved in cellular respiration, free radical defence, neurotransmitter function, connective tissue biosynthesis and cellular iron metabolism (WHO, 1998). Copper is also closely related to estrogen metabolism, and is required for women's fertility and to maintain pregnancy. Copper is also necessary for the metabolism of iron in the body and deficiency of copper may cause anaemia due to defects in iron metabolism. Also, intake of small amounts of molybdenum appears to be necessary for normal copper metabolism. On the other hand, high molybdenum intake leads to low copper absorbtion, and may induce a copper deficient anaemia (Landner and Lindeström, 1999). Copper must be in the right balance with iron and zinc, and longterm use can reduce zinc levels in the body. It is used with zinc, iron and B vitamins in the synthesis of phospholipids (long chain fatty acids) which are used in myelin formation (www.msrc.co.uk). If zink levels are too high it might lead to copper and iron deficiency. Catalyzation of the anti-inflammatory prostaglandin by cupric ions or copperchelates has been assumed to influence rheumatoid arthritis, another autoimmune disease (Landner and Lindeström, 1999).

Lower levels of zinc has been identified in serum in MS-patients compared to controls (Palm et al 1982) but other studies (Dore-Duffy et al., 1981) indicate the contrary. Stein et al, 1987 speculates that high exposure to zinc (industrial) might influence the onset of MS. Thus, Zn and other essential and non essential elements might be involved in pathogenesis of MS but the effecting factors are, as have been said above, most probably multifactorial.

3 Material

3.1 Geochemical registers

Geochemical data from soil (till), stream-water and groundwater from the Swedish Geological Survey have been used. The different available registers and their respective parameters are given in table 2. A short description of sampling density and sampling method is given below for each register as well as tables for values of the elements and parameters included in the study. For a complete description of sampling and analyzing techniques as well as data reliability on the soil and stream-water data we refer the reader to Lax (2005).

The groundwater datasets used are referred to as GW1 and are described in SSI-report 2008:15. Numbers below the detection limit have been given a value of half the detection limit. Zero values have been excluded from the study. The exact coordinate, given in the national Swedish reference system RT90 is given for each data point.

Table 2. All available data for all registers used.

	MD1	MD2	GW1	StrW
Number of analyses	38768	27695	816	37820
Alkalinitet			x	
Al			x	
Al2O3		x		x
As		x	x	x
Ba	x		x	
BaO		x		x
Ca			x	
CaO		x		x
Cl		x	x	x
Cd			x	x
Co	x	x	x	x
Cr	x	x	x	x
Cu	x	x	x	x
Fluorid			x	
Fe			x	
Fe behandlat				
Fe2O3		x		x
Hg				x
K			x	
K2O		x		x
konduktivitet			x	
Mg	x		x	
MgO		x		x
Mn	x		x	
MnO		x		x
Mo		x	x	x
Na			x	
Na2O		x		
Nb				
Ni	x	x	x	x
NH4+				
NO2				
NO3-			x	
P2O5		x		
PO43-			x	
pH			x	
Pb	x	x	x	x
Rb		x		x
S		x		x
SO42-			x	
Si			x	
Se				x
SiO2		x		x
Sr		x	x	x
TiO2		x		x
Th				
U				x
V	x	x	x	x
W				x
Yt				x
Zn	x	x	x	x
Zr		x		x

3.1.1 Groundwater

The register mainly represents private drinking water wells situated in bedrock. The total number of analyses is 816 and the sampling density is shown in figure 1. Parameters in the register are shown in table 3.

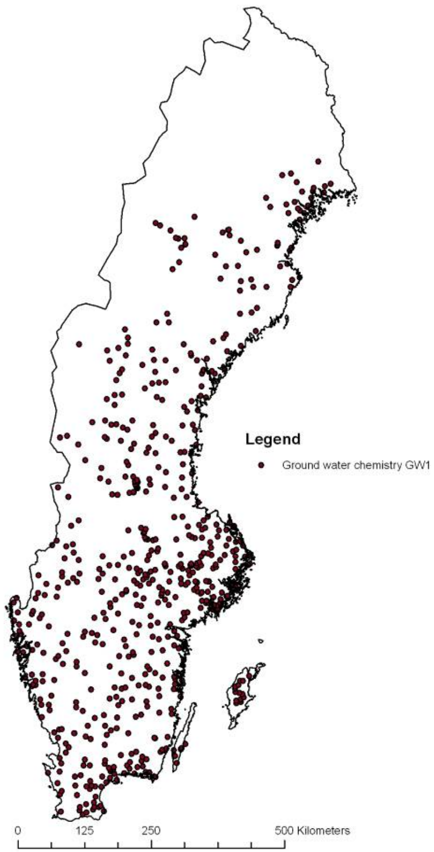

Figure 1. Sampling density for groundwater chemistry GW1. In total 816 sample points.

Table 3. Values for GW1.

	Unit	Mean	StDev	Max	Min	Median
Al	**µg/l**	35,47	176,07	4073	0,01	2,75
V	**µg/l**	0,44	0,76	8,83	0,01	0,2
Cr	**µg/l**	0,72	0,97	14,89	0,01	0,49
Fe	**µg/l**	465,18	1858,57	29300	0,3	77,78
Mn	**µg/l**	127,54	454,76	11430	0,06	39
Co	**µg/l**	0,22	1,18	30,2	0,01	0,09
Ni	**µg/l**	2,91	28,45	624,5	0,01	0,48
Cu	**µg/l**	28,66	60,80	983,27	0,05	9,75
Zn	**µg/l**	66,04	222,73	4128,21	0,05	18,61
As	**µg/l**	3,84	20,45	297,31	0,01	0,34
Sr	**µg/l**	305,28	610,66	11452,4	0,33	157,9
Mo	**µg/l**	8,93	102,88	2730	0,01	1,72
Cd	**µg/l**	0,08	0,30	5,74	0,01	0,02
Ba	**µg/l**	70,11	121,29	1183,89	0,16	31,745
Pb	**µg/l**	1,19	3,67	41,92	0,01	0,39
Ca	**mg/l**	44,33	35,30	301,15	0,25	37,17
Na	**mg/l**	37,76	56,40	404,68	0,78	15,65
Mg	**mg/l**	6,90	6,16	81,75	0,02	5,36
K	**mg/l**	3,38	4,62	53,89	0,08	2,17
pH	**mg/l**	7,37	0,65	9,9	4,85	7,43
Conductivity	**mS/m**	48,14	45,05	466,2	1,1	37,6
Alkalinity	**mg HCO3/l**	168,92	116,55	970	1	150
Sulfate	**mg/l**	24,39	55,00	930	0,22	14
Chloride	**mg/l**	33,22	89,58	973,72	0,5	12
Fluoride	**mg/l**	1,04	0,80	4,5	0,1	0,82
Nitrate	**mg/l**	3,69	9,32	75	0,001	0,22
Silica	**mg/l**	6,27	2,43	27,7	1,6	5,96
Phosphate	**mg/l**	0,07	0,20	2,6	0,01	0,02

3.1.2 Soil

Two databases of soil (till) have been used. These represent different analytical techniques (XRF and leachate) which result in different parameters and two different databases. We refer to these two datasets as MD1 and MD2. All samples are from hand-dug pits, usually 0,7-1,0 m deep. This is normally the maximum depth used, and in most cases enough to reach well into the C-horizon. The mean sampling density for till samples is one sample per 7 km² (fig. 2 and 3). For further information see Lax (2005). Mean values etc for the parameters in the register is given in table 4 and 5.

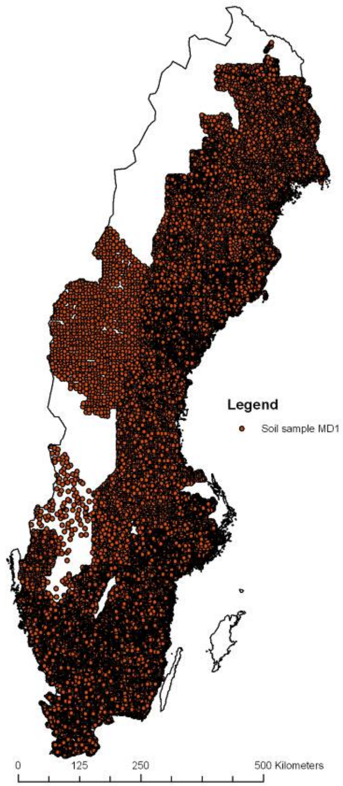

Table 4. Values for soil samples (till) MD1 (leachate). Values in mg/kg
Values marked _lev in database have been used.

	Mean	StDev	Max	Min	Median
Ba	42,03	36,52	3060	10	36
Co	5,34	3,28	63	2	5
Cr	19,77	13,92	402	1	16
Cu	15,05	14,50	1190	1	12
Mg	0,38	0,23	4,57	0,001	0,33
Mn	257,18	221,22	9460	10	201
Ni	11,37	8,98	283	2	9
Pb	11,35	13,82	1080	7	8
V	28,44	14,47	504	6	26
Zn	37,04	28,44	2197	5	32

Figure 2. Sampling density for soil samples MD1. In total 38768 sample points.

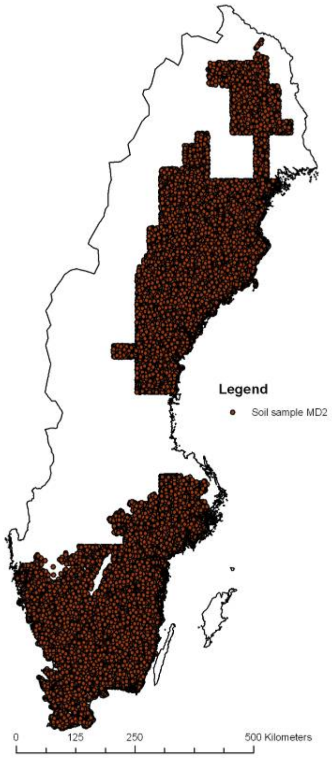

Figure 3. Sampling density for soil samples MD2. In total 27695 sample points.

Table 5. Values for soil samples (till) MD2 (XRF-method). Values in mg/kg

	Mean	StDev	Max	Min	Median
Al2O3	13,6	1,7	33,1	1,3	13,6
As_	9,0	7,1	229	5	5
BaO	0,1	0,0	0,847	0,019	0,06
CaO	2,3	1,6	54,98	0,09	2,16
Cl	139,8	253,1	5193	25	68
Co	21,0	6,6	95	2,5	20
Cr	52,6	27,0	604	6	48
Cu	16,8	14,1	672	1	14
Fe2O3	3,9	1,2	15,38	0,49	3,71
K2O	2,9	0,5	6,77	0,37	2,86
MnO	0,1	0,0	1,124	0,009	0,059
Mo	1,6	4,6	659	1	1
Na2O	2,7	0,7	5,2	0,1	2,7
Ni	18,7	11,4	204	2	16
P2O5	0,2	0,1	1,969	0,01	0,232
Pb	24,0	15,0	1191	5	22
Rb	89,1	23,9	408	5	85
S	211,8	316,1	29494	25	147
SIO2T	72,0	3,9	89,2	41,1	72,1
Sr	175,8	55,4	706	21	173
TIO2	0,739	0,185	2,596	0,069	0,736
V	65,0	30,3	1562	5	59
Zn	54,3	31,0	2165	8	50
Zr	476,0	148,1	2243	56	445

3.1.3 Stream-water

The stream-water register contains data from aquatic plants (roots) and bryophytes from small first- or second order streams, thus reflecting drainage basins of a few square kilometres. Most samples are composed of roots of sedge (*Carex L.*), willow moss (*Fontinalis antipyretica*) and roots of meadowsweet (*Filipendula ulmaria*). The mean sampling density in the regional sampling programme is approximately one sample per 7 km^2 (fig. 4) but sampling density for cadmium, selen and mercury is not as dense as for other parameters. The exchange of metals between the water and the roots is a slow process whereby the influence of seasonal variations is of minor importance. The metal concentration in the stream water reflects the chemical composition of the surrounding bedrock and soils but is also influenced by anthropogenic pollution. For more information about the register, see Lax (2005). Data included in the register is shown in table 6.

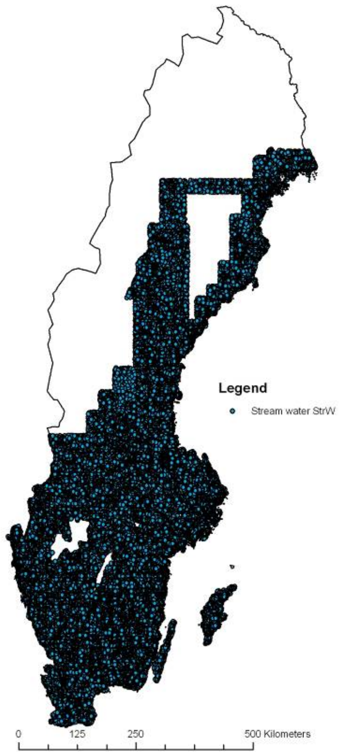

Figure 4. Sampling density and area covered for stream water data. In total 37820 sample points.

Table 6. Values for stream-water.

	Mean	Stdev	Max	Min	Median
Al2O3	2,02	1,66	18,04	0,007	1,538
As	16,44	56,29	3584,18	0,01	6,37
BaO	0,03	0,03	0,621	0,001	0,019
CaO	1,20	1,10	38,61	0,05	0,98
Cl	1333,46	1903,78	54374	1	808
Co	24,07	35,52	1580,8	0,2	14,7
Cr	11,67	51,37	9366,2	0,1	8,5
Cu	15,20	58,19	8830,1	0,2	11,1
Fe2O3	4,13	3,69	54,02	0,04	3,03
K2O	1,01	0,64	17,52	0,01	0,89
MgO	0,42	0,27	9,38	0,02	0,35
MnO	0,61	1,27	25,558	0,001	0,144
Mo	3,60	6,44	362,29	0,01	2,07
Ni	11,62	24,25	2029,7	0,1	6,8
Pb	27,46	66,18	8338,1	0,3	18,9
P2O5	0,37	0,27	4,24	0,02	0,29
Rb	30,59	18,40	499,2	0,15	27,2
S	3376,95	1843,81	34645	10	2956
SIO2	9,52	8,98	72,12	0,03	6,7
Sr	53,70	34,09	841,7	0,5	47,9
TIO2	0,09	0,09	1,428	0,0003	0,0642
U	4,64	10,07	870,32	0,01	2,25
V	31,32	22,51	486,1	0,2	26,8
W	1,89	3,72	351	0,1	1,4
Yt	19,87	21,53	683,6	0,1	14,5
Zn	122,04	240,62	18802,4	0,7	71,3
Zr	48,12	48,93	962,4	0,1	32,8
Cd	1,300	2,459	74,55	0,005	0,72
Hg	0,0598	0,0467	1,415	0,0005	0,049
Se	0,499	0,717	45	0,005	0,36

3.2 MS-register

We used the Swedish MS-register, which includes almost all MS-patients in Sweden. The register holds information on birth year, birth place, where the patient live and if the MS patient was born abroad. The register also holds information on when MS was first diagnosed (eg. the onset age, the age when diagnosed), if the patient is a woman or a man, and year of death if the patient is deceased. The register contains 9107 persons in total of which 6454 are woman and 2653 are men. Full information is not available on every record. 351 records lack information on where the patient live. 591 records lack information on year of onset. When all relevant criteria are combined; living place, birth year, year of onset, there are 8076 MS patients where 5759 are women and 2317 are men. A diagram showing the age distribution year 2007 of all MS-patients in the register sub divided into males and females is shown in figure 5. A diagram showing the distribution of the onset age of MS for all patients as well as for women and men separately is showed in figure 6. The onset age for women is normally between the age of 20 and 40 while for men displays a narrower span between 25 and 35.

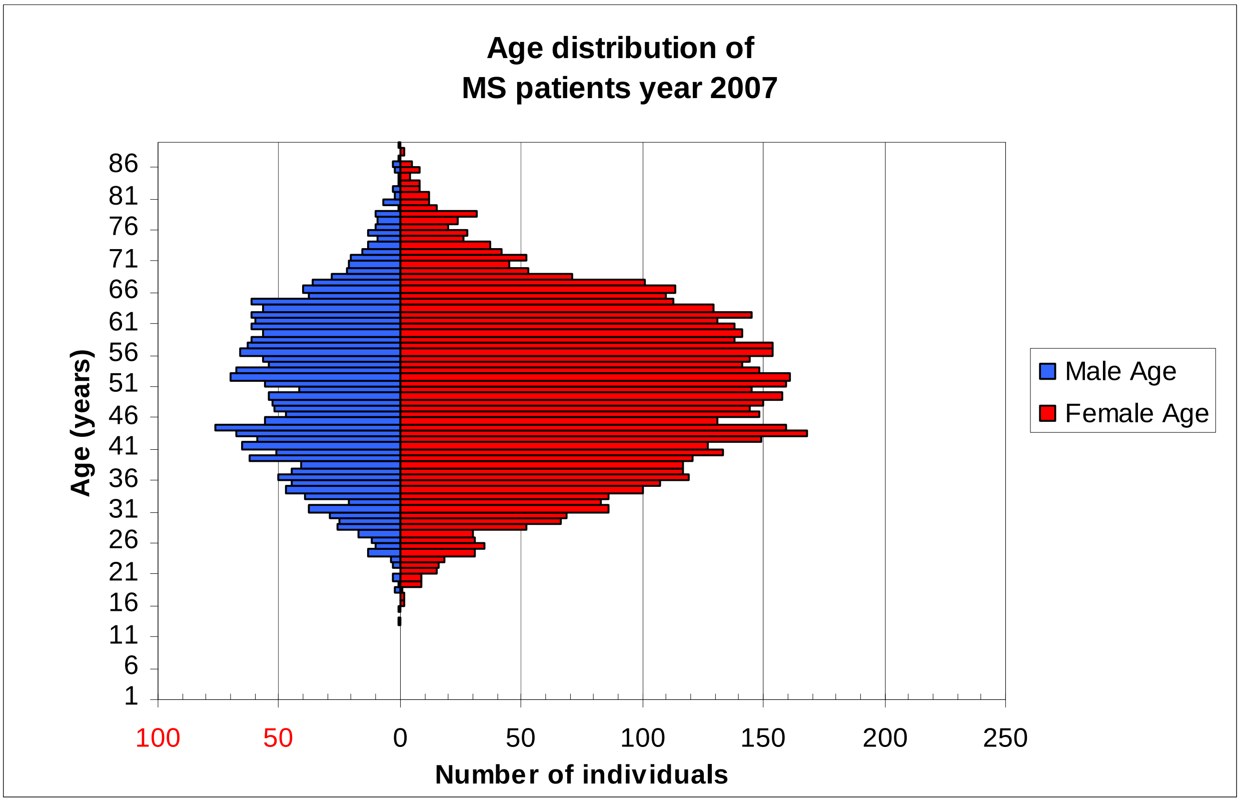

Figure 5 Age distribution of MS patients in Sweden.

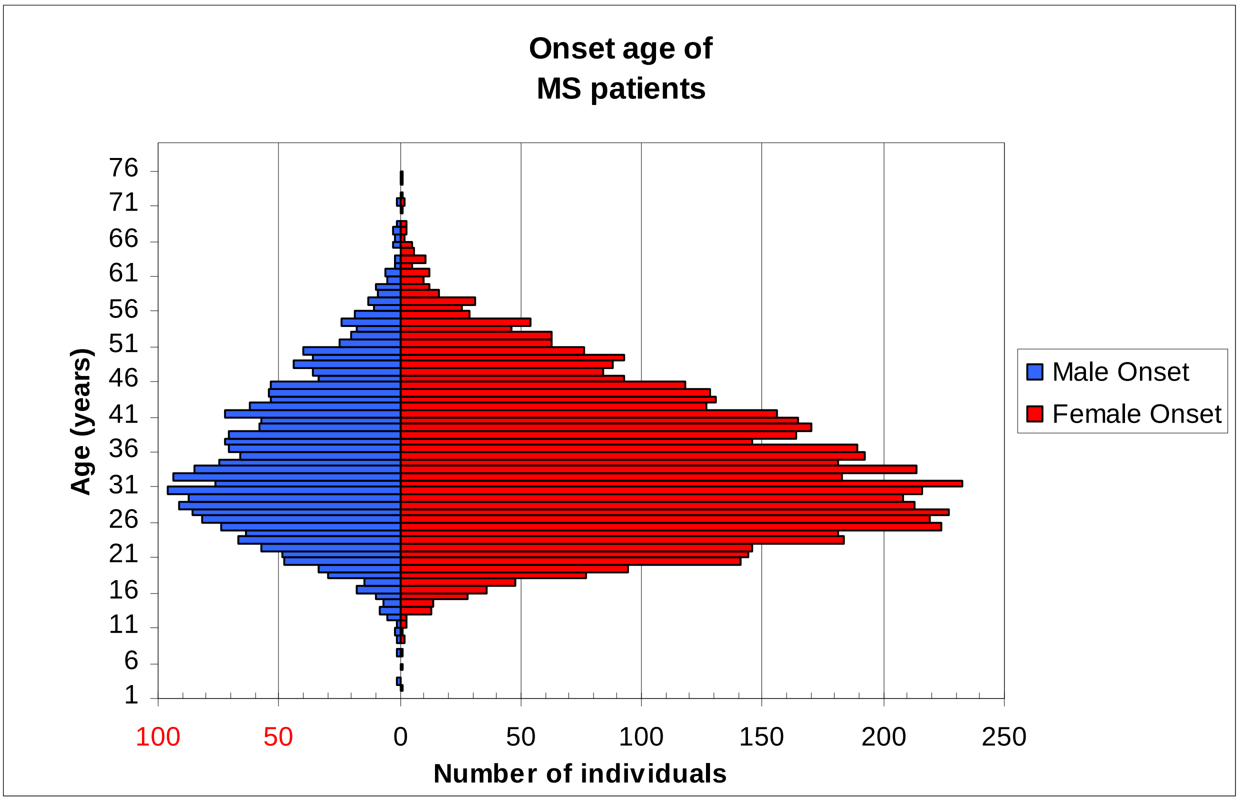

Figure 6 Distribution of onset age of MS for men and woman in Sweden.

3.3 Census data

Spatially distributed census data, valid for December 2005 and compiled by Official Statistics of Sweden, Statistics Sweden (SCB), were used in this study. The number of persons in the census data is grouped into post code areas (postal-CA), the smallest available administrative unit with a geographical context used in Sweden. The total number of postal-CA´s is 9061 but 99 of them lacks inhabitants, commercial or industrial areas etc. The size, in number of inhabitants and in geographical extent is extremely variable and is ranging from 0 to 6672 inhabitants and from 855 square metres to 12058 square kilometres in size, with an average size of 47.4 square kilometres. All postal-CA´s can be grouped into 1622 post districts. The number of postal-CA´s within a post district is ranging from 312 in Stockholm to one single postal-CA. This take place when a postal-CA makes up the whole area of a post district, which occurs at 951 places in Sweden.

Post code areas are constructed in such a way that the number of inhabitants is approximately equal to 1000 people. This has the effect that in areas with dense population, the size of a post code area will be small and vice versa for areas with sparse number of inhabitants. As a result, post code areas are not very well suited as geographical units for the distribution of natural occurring geochemical elements but is the smallest common geographical unit that can be used to combine geochemical data and diagnosis of MS on spatial basis. In this study post code areas and post districts are used as spatial reference.

4 Methods

4.1 Spatial Distribution

All datasets were imported into a database with built in support for spatial quarries.
The coordinate pair of a sample point in the geochemical dataset was used to calculate the spatial belongings to a specific postal-CA for each sample point. The calculated postal-CA code was stored as an attribute for each sample point. In MS-patient data all addresses were given as postal-CA codes. The number of MS patients were summarised over postal-CA codes giving the total number of patients within a postal-CA and thus within a post districts. Using postal-CA codes, geochemical data, MS- data and census data is combined on a spatial base with postal-CA´s as the smallest spatial common denominator. The relationship for the combination of datasets is shown in figure 7 below.

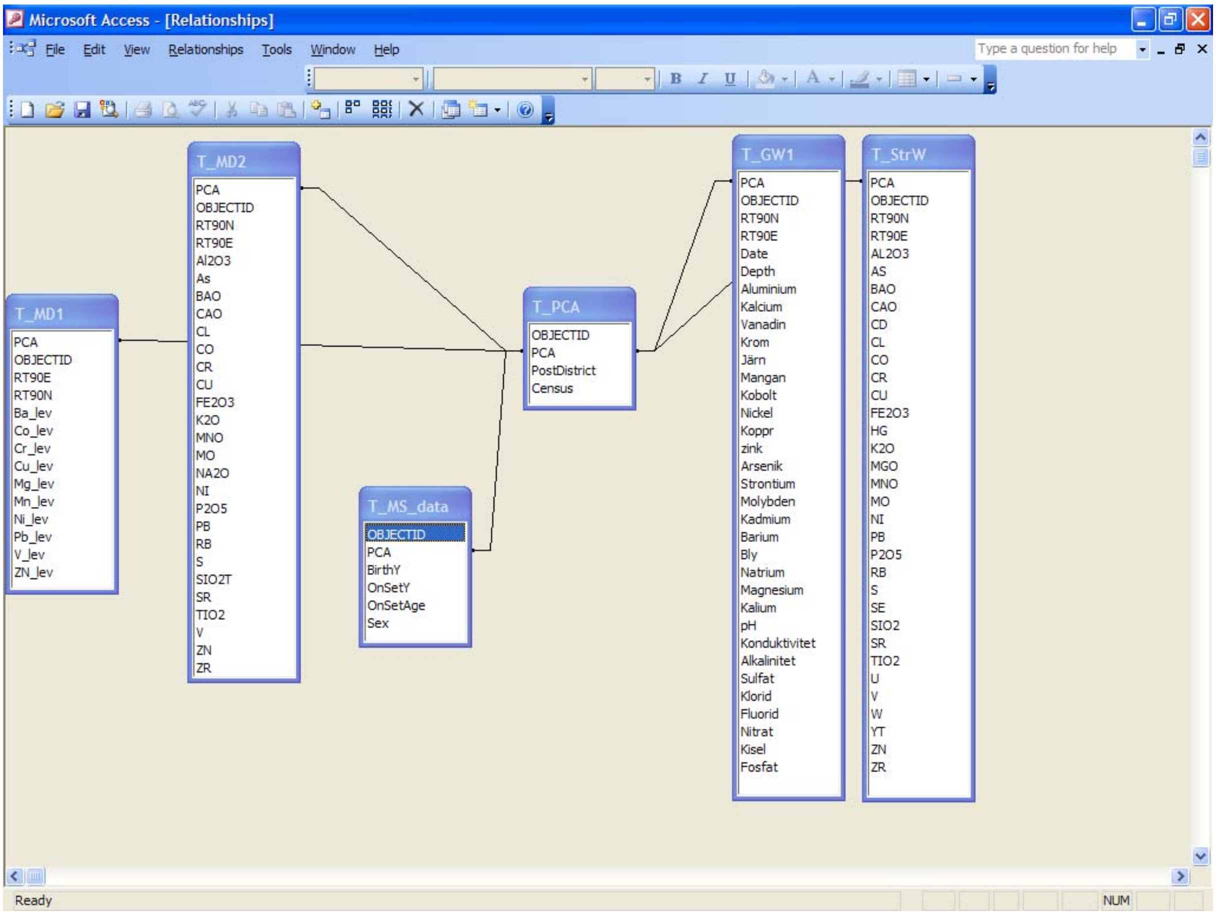

Figure 7. Database relationships view between datasets. All datasets are linked using Post Code Area (postal-CA) as the key.

For the purpose of this study, prevalence is defined as the ratio of the number of MS patients in relation to the total number of inhabitants within a postal-CA or a post district. Prevalence is scaled by a factor of 1000 to reduce the number of decimals needed to maintain a high degree of precision. Prevalence measures are calculated and saved as an attribute for each post code area and for post districts respectively.

A better measure of the occurrence of MS would have been the incidence (i.e. the number of new diagnoses that occur in a given time period = the rate of the disease) but since MS has a very unspecific onset diagnosis the incidence is a difficult measure to calculate for this disease.

Since the size distribution of post code areas vary considerable over the country, prevalence measures gave extreme values in areas with few inhabitants and occurrence of MS. Furthermore, visual interpretation of prevalence measures, as shown in maps below, is strongly biased towards large post code areas, masking the variation of prevalence measures of small areas. This effect is striking in larger cities always having a large number of small post code areas (fig. 8).

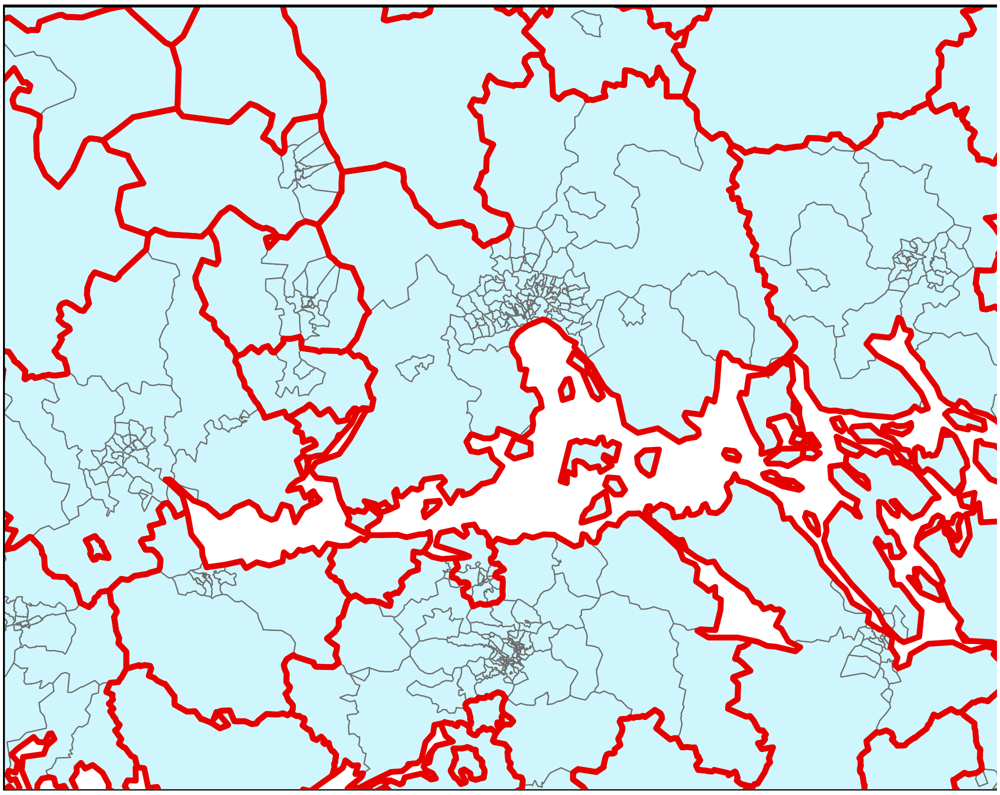

Figure 8. Post code areas, grey line and post districts, red line, for Västerås East of Stockholm. White areas is made up by lakes.

4.2 MVA-methods

Combination of the registers was evaluated with multivariate analysis (MVA). Typical examples of MVA methods are principal component analysis (PCA) (Martens and Naes, 1989; Wold et al, 1987) and partial least squares (PLS) (Martens and Naes, 1989; Geladi and Kowalski, 1986). Both techniques reduce the multidimensional data set to lower dimensions by calculating so-called principal components (PCs) that describe the data. A PCA model is based on the X-block (i.e. the frequencies) and calculated in such a way that it describes as much variance as possible in the data, whilst a PLS model also takes the correlation to the response(s) of interest into account. Results from PLS and PCA are often interpreted in score plots and loading plots. Score plots show how the samples are distributed and loading plots display the relationships between the variables. Figure 9 below shows a geometric interpretation of PLS.

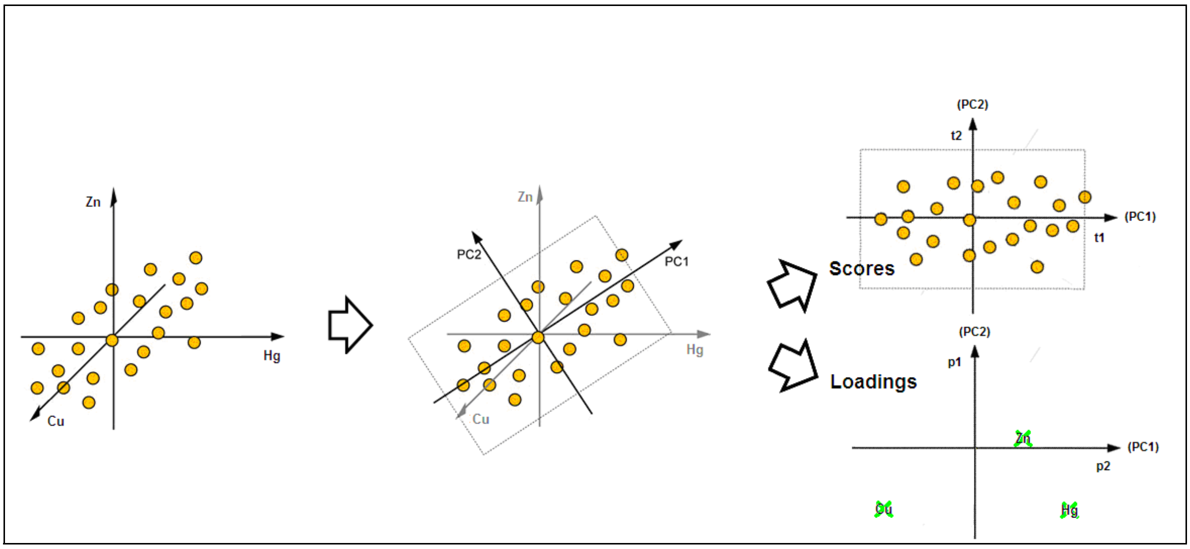

Figure 9. a. Each sample has a value for each concentration (eg. Fe, Cd, NO_3), giving it a coordinate in the n-dimensional space (n = number of variables (frequencies)). Each sample also has a corresponding prevalence value. b. A number of principal components (PCs) are placed in the n-dimensional space in such a way that they describe the data as good as possible. c. The score plot shows the projection of the samples on the PC plane and the loading plot shows the influence of each variable on the PCs.

Data have been analyzed with multivariate program SIMCA-P+11.5. Principal component analysis was performed on stream water (BV/StrW), soil (MD1 and MD2) and groundwater (GW1) of the separate tables and altogether in a large multivariate matrix (all registers). Note that these tables where randomly split in half with the constraint that post code areas/postal areas should be distributed fairly evenly over the entire country. A set of untouched data was spared throughout the work, aimed for final verification of results only. Validation of the PCA model was done by cross validation, using 8 random segments. All variables (both concentrations and prevalence values) were power or logarithm transformed to reduce skewness of the distributions before centering and standardizing them to unit variance. Partial least squares, PLS models, were built to connect stream-water data, groundwater data and soil-data with prevalence of MS.

4.3 Delimitations and assumptions

This study is subject to the following delimitations and assumptions:

- The MS-registry is not complete, thus, not all patients with MS in Sweden are recorded in this study.
- The 351 MS patients in the register from whom there is no information about where they live have been excluded in this study. Also, deceased people have been excluded since they would give a false indication of the prevalence (i.e. prevalence would be accumulated over time).
- We have neglected that people move during the course of the disease. Neither is such data kept in the MS-registry.
- We have considered MS as one disease and not divided patients according to different diagnoses.
- Areas where no geochemical data have been available for a post districts have been excluded.

- In this study we have used the geographical distribution of post code areas and the population density from year 2005. We have neglected that postcode areas might have changed their geographic distribution over time and consequently also the population density.
- We have assumed that the geochemistry data in the different medias are constant over time. This is false, e.g. due to seasonal changes, at least for water. However, we have assumed that this change is small and also averaged over time and can be neglected in this study. Another source of error is the uncertainty of the analytical methods, but compared to other sources of error in this study it is assumed to be of minor importance.
- The available groundwater data are mainly taken from drinking water wells where the water has passed through pipes and technical devices that are made of copper and /or other metals. Cleaning devices such as filters etc can also be present. This will influence the content of metals in the drinking water to no longer be reflecting true background levels. It is nevertheless a true value for exposure. For this reason we have not excluded copper from the study. We have instead considered the elevated levels of copper due to water passing through pipes to be more or less constant over the country and the natural background level to be varying. This might be false since different distances to wells (i.e. longer/shorter length of pipe etc) and quality of pipe will differ between wells.
- We have considered the population in Sweden to be genetically homogenous, i.e. genetics is not included in this study.

5 Results and discussion

5.1 Prevalence of MS

The calculations and statistical treatment yielded some information about the distribution of patients in the MS-register. The results show that there is a difference in prevalence in different areas in Sweden (fig. 10). There are 639 out of 1622 (39 %) post districts with no MS data at all. The number of male and female MS-patients is grouped differently to maximize the visual interpretation of spatial distribution but the relative difference is still the same. There seems to be higher prevalence in the county of Västerbotten and clusters around the larger cities. This might relate to population being denser in the areas around cities. It could also be a consequence of the location of the larger hospitals. The map is somewhat difficult to interpret due to the size of the postal districts where large areas with high prevalence stand out very clearly, risking misinterpretation. Maps showing the distribution of MS for men and women separately are shown in figure 11 and 12 (number of cases, not prevalence). From these maps it is clear that high numbers of MS-patients are observed close to larger cities and that the pattern for women is more widespread over the country. Prevalence for women and men could not be calculated due to lack of census data separated on men and women.

The map in figure 10 would have been easier to interpret if the prevalence would have been calculated for other units than postal districts. A preferable unit for this study would be catchment areas or areas with approximately the same size, but without loosing precision, i.e. they can not be as large as counties. To calculate the prevalence for catchment areas population density for such division is necessary.

Sun exposure was not included in the analysis since no north – south gradient was found in the data.
Similarly, no east – west gradient was found.

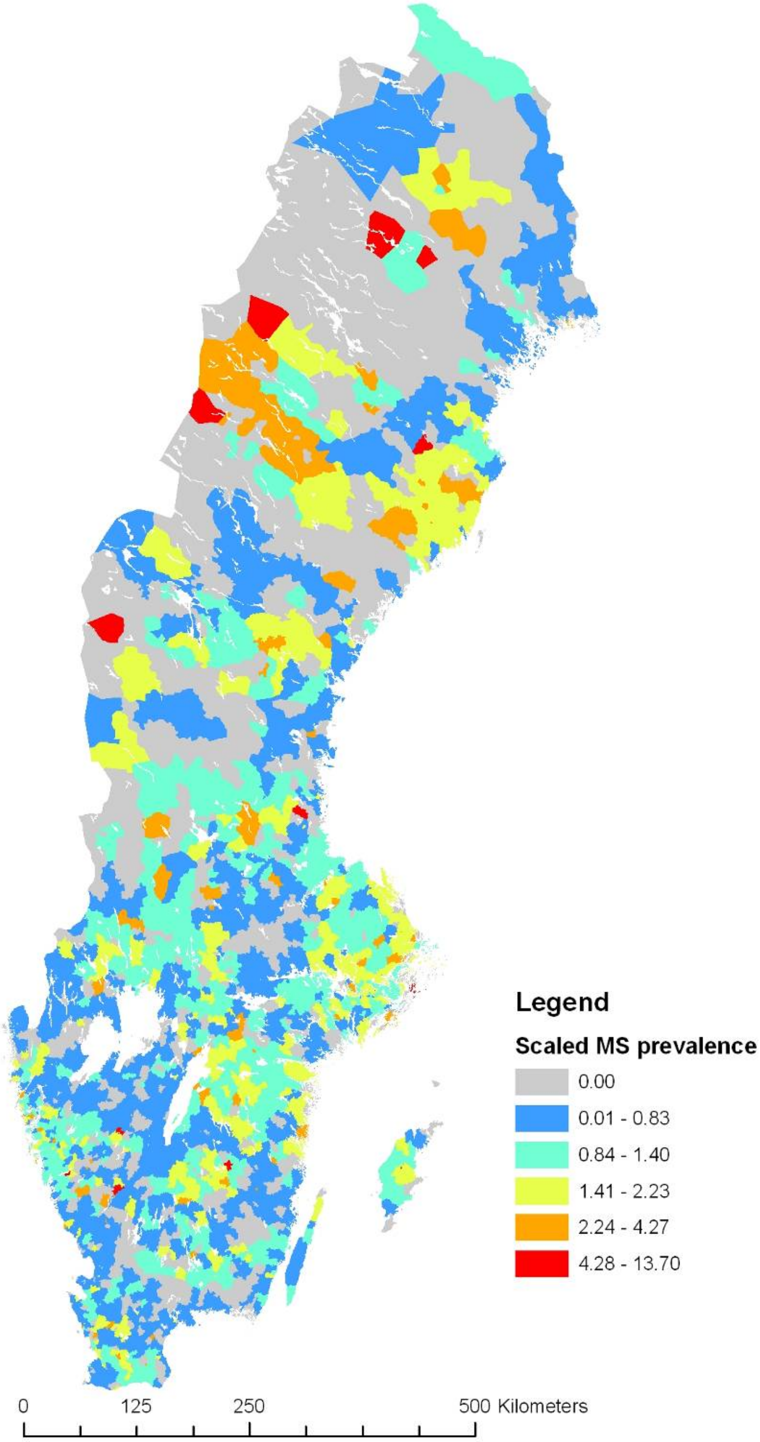

Figure 10. Prevalence, the ratio between the sum of MS patients within a post district and the number of inhabitants for the same district scaled by a factor of 1000.

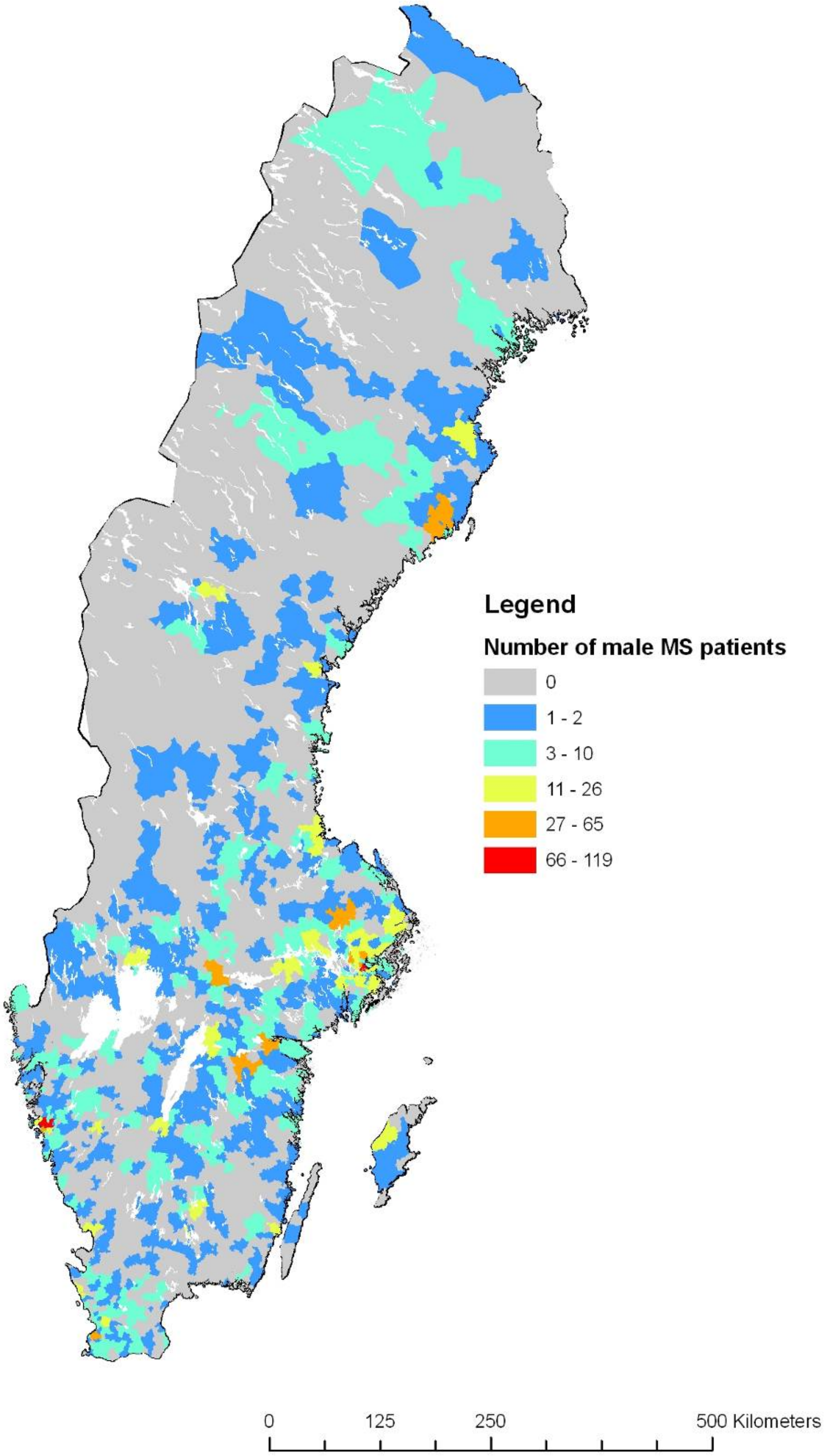

Figure 11. Map showing number of MS patients (men) in postal districts. Grouped by natural breaks in data.

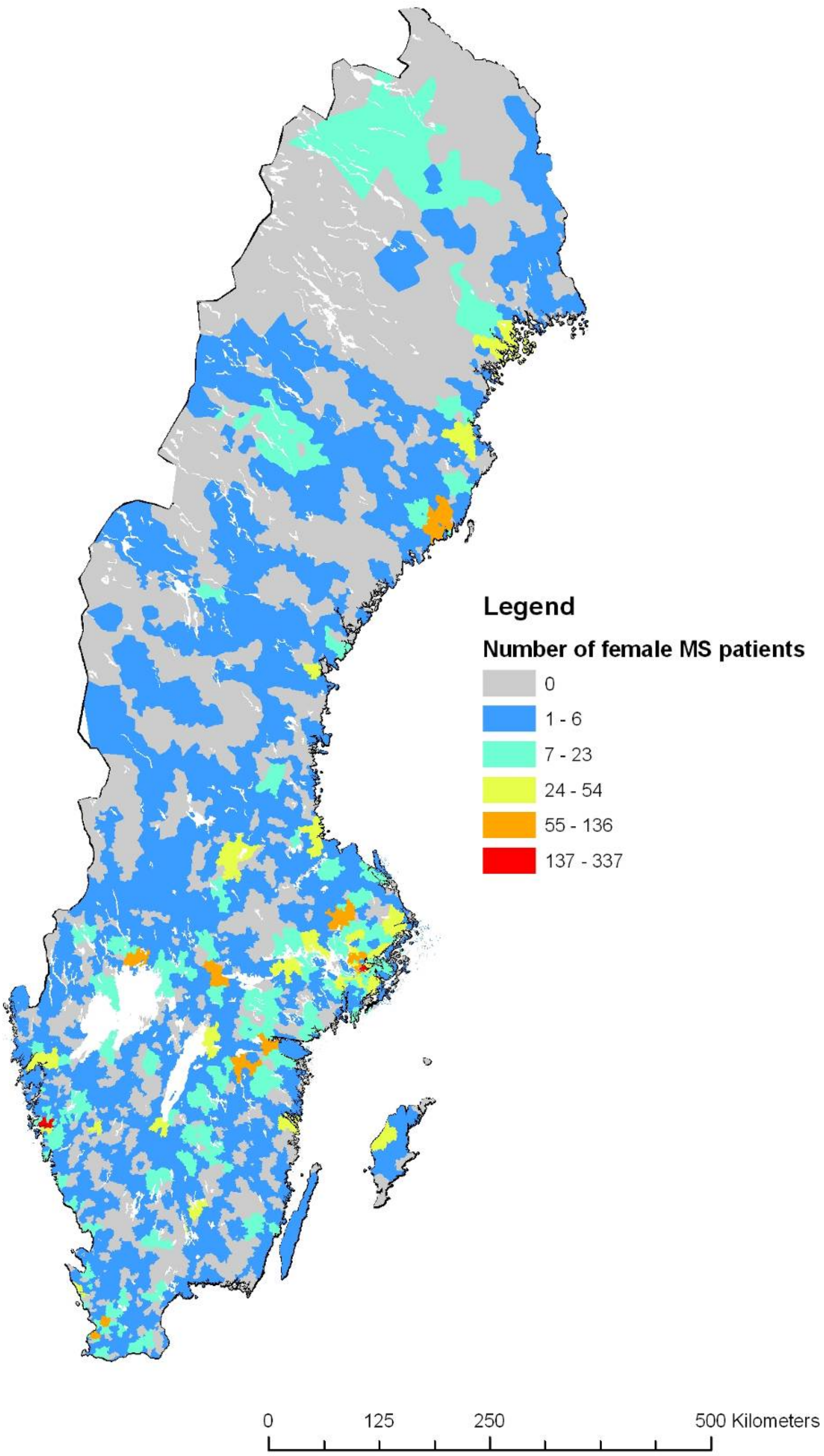

Figure 12. Map showing the number of patients (women) in postal districts. Grouped by natural breaks in data.

5.2 Combination of registers

5.2.1 Univariate overview of Zn in different registers

First a univariate comparative analysis of Zn in the three registers was done by presenting data on maps. Three maps showing the average values of Zn over post districts in the three different medias is shown in figure 12. They pinpoint the problems that arise with averaging over post districts and areas falling out of the analyse as a consequence of the sparse sampling net, this is particularly obvious for groundwater. (i.e. if a post district does not have any sampling point it has not been included in this study). Also, comparison of the map of soil samples (left in figure 13) with the sampling density map of MD1 (fig. 2) shows that some areas have got a mean value where there are no sampling points but still a post district available. In other words, some post districts are large and have few sampling points (low density of sampling points) and will be given an average value that is maybe not representative for the whole area.

The maps of soil and stream-water show some similarity of higher zinc levels in the Skåne region (southern tip of Sweden), around lake Mälaren and Siljan but the picture is not very convincing. The maps in fact demonstrate that the registers represent different medias and different exposure pathways. The soil samples shouldn't be affected by air deposition while stream water samples are both a consequence of the combined natural background and anthropogenic influence as well as the variable uptake of elements (in this case Zn) in the investigated specie. The coverage in the ground water map is too sparse to be able to conclude on any geographical variation of Zn.

5.2.2 Univariate analysis of prevalence of MS and Zn

The prevalence of MS was checked against Zn solely but no correlation was found. Average values over both postal-CAs and post districts in all three registers were used. One way to evaluate if this correlation is random or true would be to fit a linear model between Zn and the prevalence. In this case it was found that the distributions were too skewed for at least a square fit to be reliable. Therefore both variables were added by one before taking the log10 transform on them. Figure 14 shows the distribution of prevalence and Zn data. A linear fit is applied on the transformed Zn and prevalence data in figure 15. Confidence intervals were calculated for the intercept and slope. The slope (m) was found to be significant on 95% level but the intercept (k) was not found to be significant at the 95% level, for more information see table 7. This indicates that the prevalence increases with the Zn concentration. However, this result is not that conclusive since the intercept is not significant, e.g. we can not give any indication when log(1+Prevalence) is zero.

Soil (MD1) Streamwater Groundwater

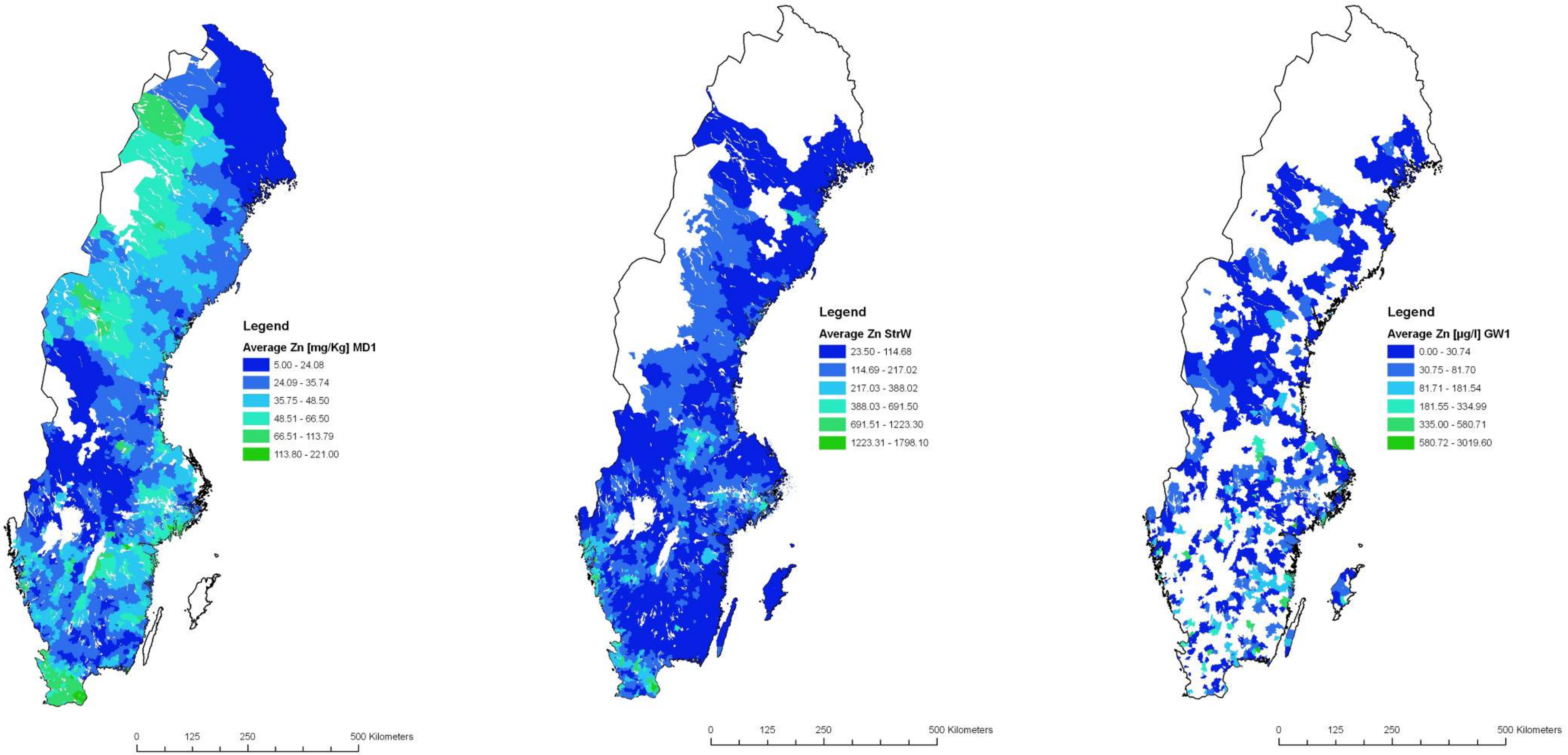

Figure 13. Maps showing the mean values for post districts in the different media: soil (MD1), streamwater (StrW) and groundwater (GW1).

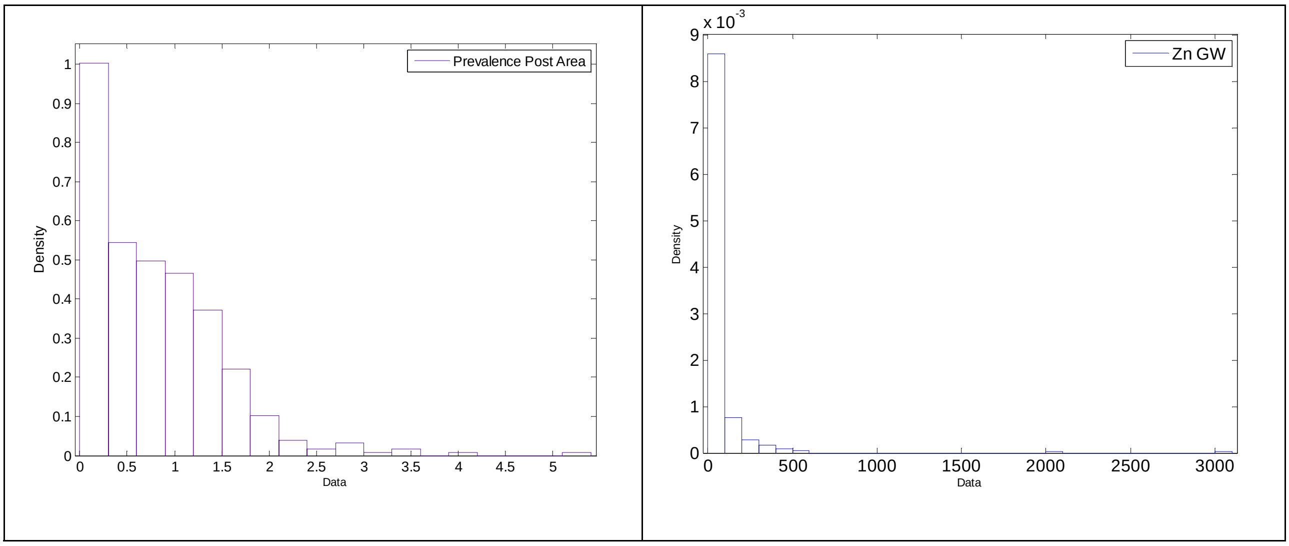

Figure 14. Distribution of prevalence data (left) and distribution of Zn-data in groundwater (right).

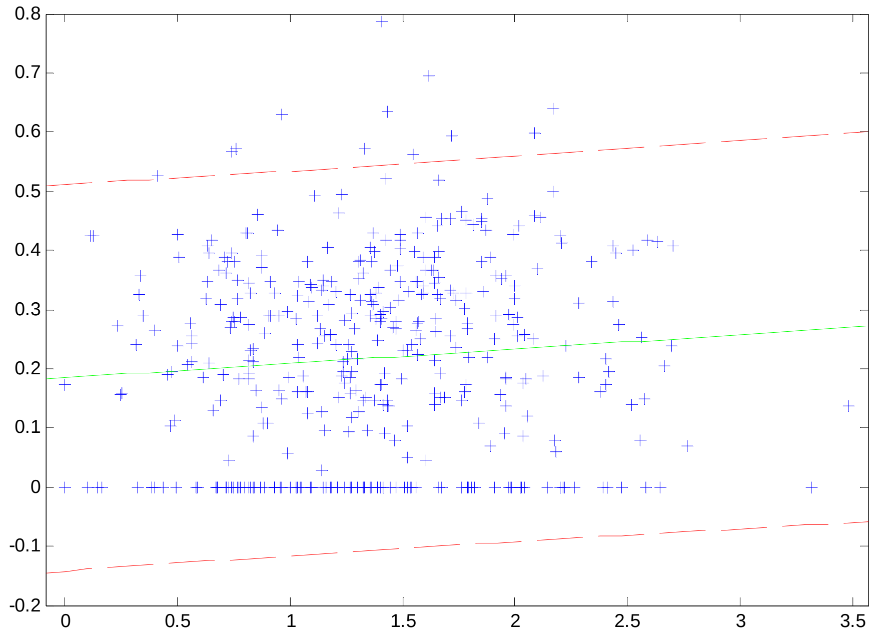

Figure 15. Linear fit of prevalence and zinc in groundwater.

Table 7.

	Parameter value	Lower CI	Upper CI
k	0.0242	-0.0026	0.0510
m	0.1847	0.1451	0.2242

5.2.3 Multivariate analysis MS and geochemistry

Unfortunately, none of the PLS-models explain more than a few percent of the total variance, therefore showing results from them would be questionable. Registers were analyzed both separately and together, but none of these models increased the degree to which variance was explained. Analyses were made both with all values included and with extreme values excluded.

The use of post districts instead of post code areas will result in loss of geographic specificity but give a more stabile average value.

Further work would be needed analysing these models if any further conclusion is to be drawn.

5.2.4 Multivariate analysis of geochemical registers

The soil, stream-water and groundwater datasets were analysed as PCA with all parameters included and plotted separately. Outliers have been excluded. Analyses were made both as log-values and as normal values but positions of the loadings remain the same. The reason for using log-values is a) that the distributions are skewed and b) when exporting data from the GIS-tool missing values internal represented as null are exported as zeros. By logarithm transforming geo data the two objects were met, most of the distribution became more normal and the logarithm of zero yields "not and number", NaN (handled as missing data in MVA software). Loading (the weight that turns that original data into the PCA-scores) graphs showing the main two variations in for three registers are shown in fig. 16, 17 and 18.

It was also investigated if there was any difference between loadings if datasets of average values over postal-CAs or the raw data were used. Small variations could be seen but overall the loadings appeared the same.

The plots show that there is a difference between the registers and that elements and substances are behaving differently in different medias. They also show that the mutual correlation between elements is similar for some substances even if the media is different. For example Lead (Pb) and arsenic (As) appear close to each other in the soil and stream-water plot while they are far away in the groundwater plot. In the groundwater-plot (fig. 18) the loadings show classical correlation between sodium and chloride, calcium and magnesium and the heavier metals cluster together in the upper right corner.

The Zn, Pb, Cr, Cu and Ni- values from stream-water (StrW) and soil (MD1) databases were also analysed in the same model to investigate if positions would overlap for the elements. These two datasets have approximately the same sampling density and coverage. Results are shown in figure 19. As can be seen, the registers separate. This indicates that there is more variance between these two registers than between parameters in the register and that parameters in the different registers could not be used as proxies for each other.

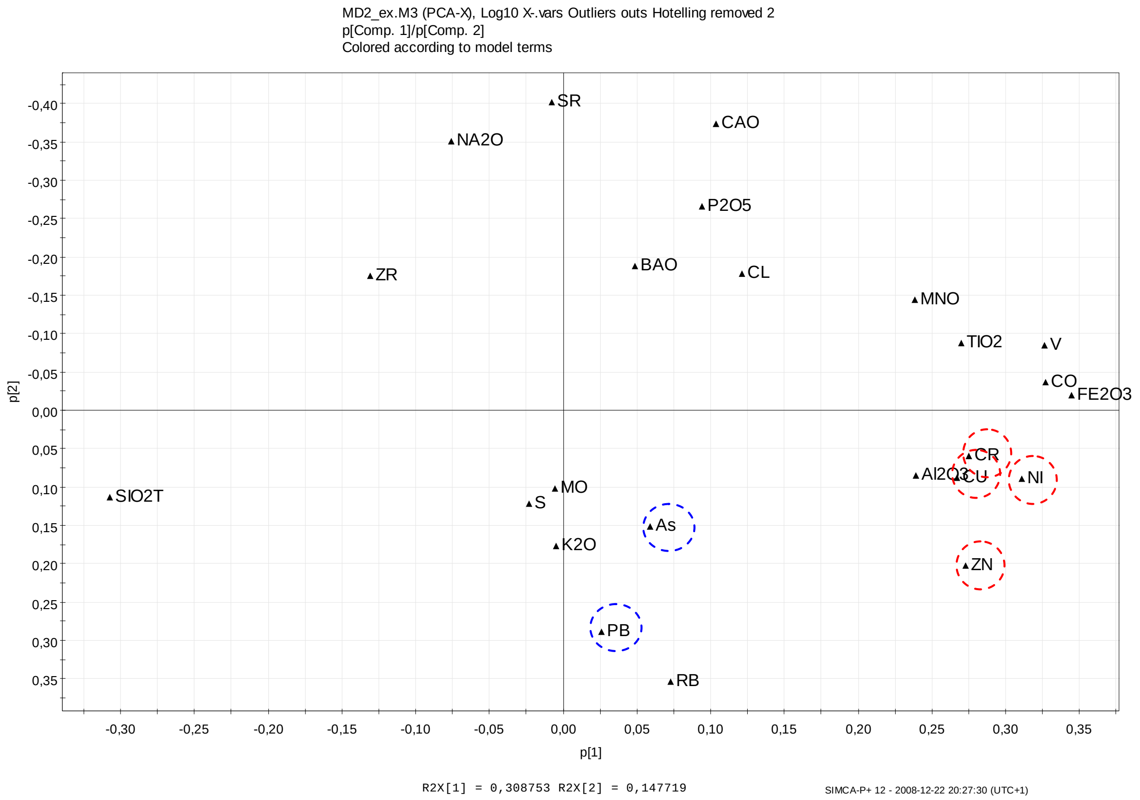

Figure 16. Loadingplot for parameters in the soil register (MD2) (exact coordinates, outliers removed, y-axis turned). This two component PCA-model explains 45.6 % of the total variance of this data set of which 30.8 % is explained by the first component.

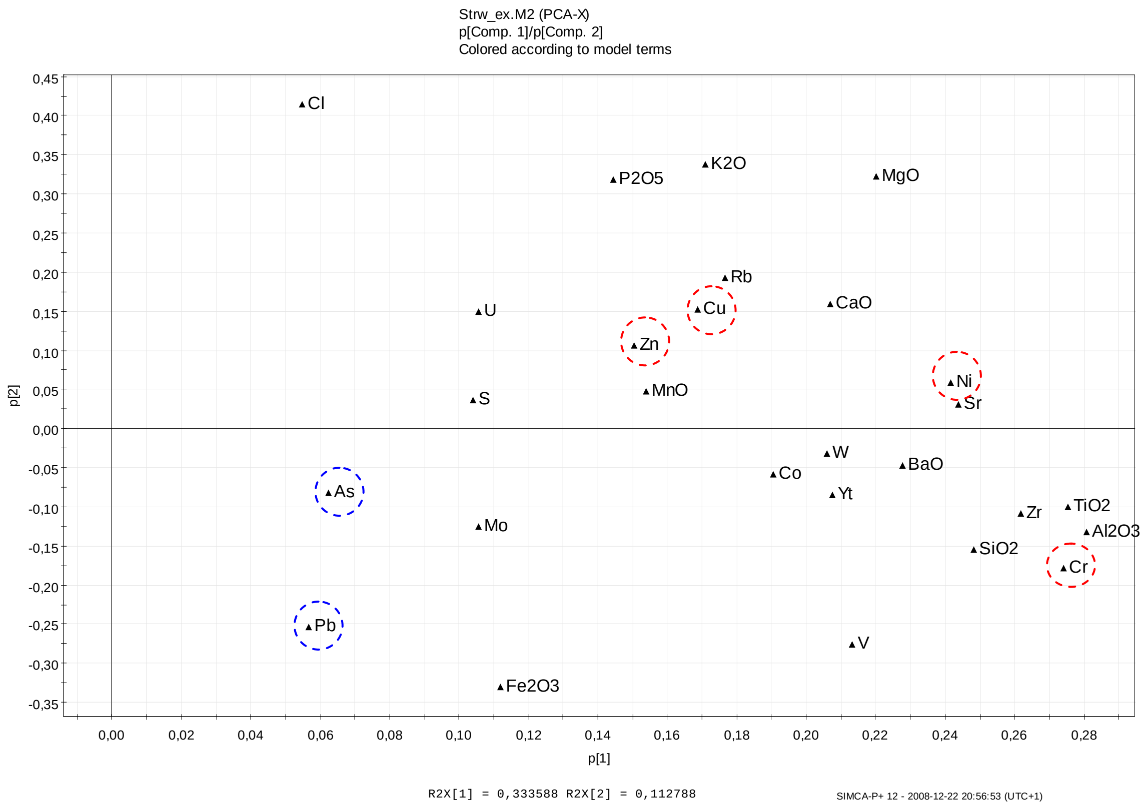

Figure 17. Loadingplot for parameters in stream-water register (exact coordinates, outliers removed). This two component PCA-model explains 44.6 % of the total variance of this data set of which 33.3 % is explained by the first component.

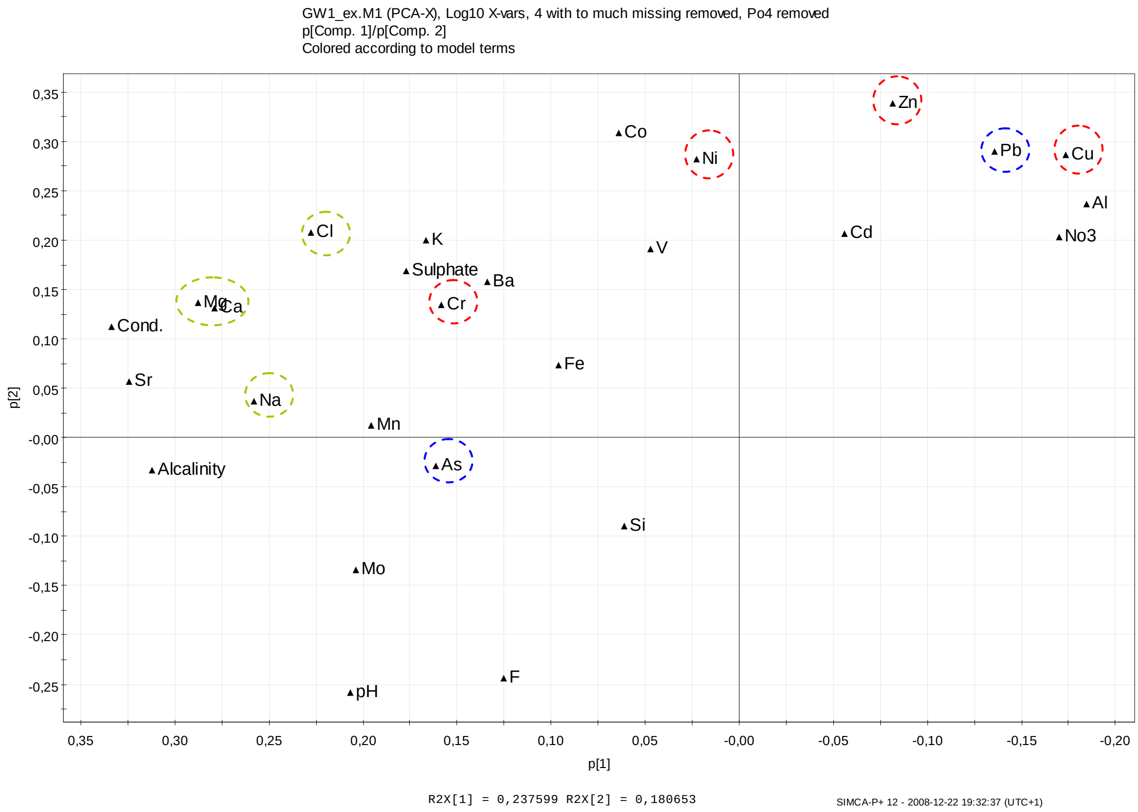

Figure 18. Loadingplot for parameters in the ground-water register, phosphate and four outliers removed. This two component PCA-model explains 41.8 % of the total variance of this data set of which 23.8 % is explained by the first component.

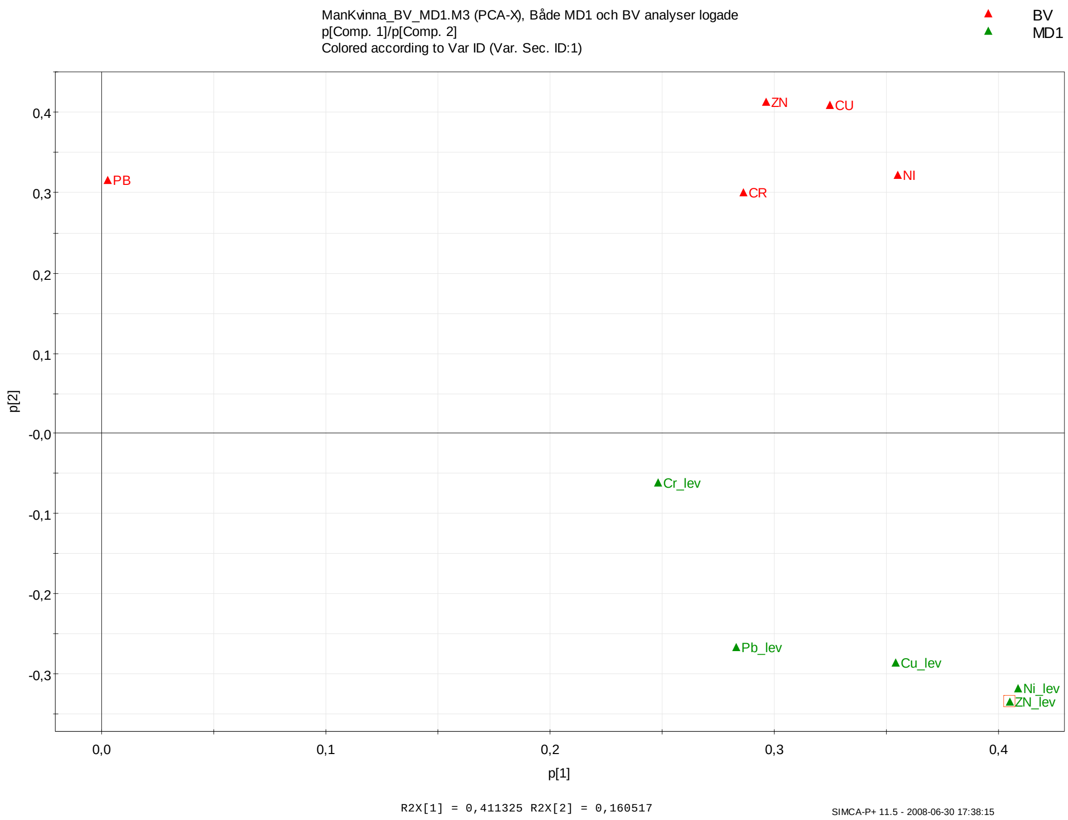

Figure 19. Loadingplot for stream-water and soil (MD1) data (Zn, Pb, Cu, Ni, Cr). Red symbols = soil and green symbols = stream-water. (Mean values over postal districts, 822 districts)

6 Conclusions and recommendations

Based on the above analysis, we make the following conclusions and recommendations:

- The analyzed data on the distribution of MS-patients indicate that a geographical pattern could be found with higher prevalence of MS in the county of Västerbotten and clusters around larges cities. However, visual interpretation of prevalence measures is strongly biased towards large post code areas, masking the variation of prevalence measures of small areas. This effect is striking in larger cities always having a large number of small post code areas.

- No north –south or east-west gradient of the prevalence was found.

- No significant and verified correlation with prevalence of MS and geochemistry of soil, stream-water and groundwater in Sweden was found. This does not mean that no connection is possible but that correlations could not be found with the data, methods and models used in this project.

- To combine patient data with exposure data with a geographical variation, the administrative division (i.e. in parishes, post code areas etc) are less appropriate. Divisions with respect to natural (geographical) borders such as catchment areas are more constructive for epidemiological purposes when a geographic component is of interest.

- The density of the patient data and the exposure data is of crucial importance. Moreover, there must be a variation in the exposure data large enough to result in a difference between areas.

- It is recommended that also the areas where no or low level of disease is found to be included in epidemiological studies, since high or low levels of different elements could exist even in these areas.

- The use of average values over districts is problematic. A high density of sampling in an area does not necessarily mean that the calculated mean value is representative for the whole area. How well an average value for a district describes the actual value depends both on the natural variability of substances in the media as well as the sampling density (i.e. high variance but many samples could give the same average value as low variance and few samples).

- The sampling of different areas and medias is not performed during the same season and time span. This will result in seasonal variations to be built in to the average values.

- The behaviour of elements in different medias like stream-water and soil is similar for some parameters but generally the variance is larger between these two registers than within.

- To be able to calculate prevalence for men and women separately in post code areas, data of male and female distribution in the population for post code areas (or the unit used) is needed. This is especially important for MS since the disease is not evenly distributed between men and women.

- The lack of information about where the patients lived before their MS debuted is a major source of error when determining the connection between elements and prevalence. In addition there is also likely a dose and time relationship, which is cumulative exposure to the elements studied.

- Exposure from other media, for example air, could easily be included in further studies.

7 Acknowledgement

Thanks to the SGU-staff Olle Selinus, Kaj Lax and Britt-Marie Ek for valuable expertise and help with the geochemical registers. We would also like to thank Leszek Stawiarz and Oluf Andersson for help with the MS-data and Ann Marie Landtblom for helpful comments on MS-research.

8 References

8.1 Literature

Ahmed, S., A. Gyllenberg, et al. (2006). Major histocompatibility complex, class II, transactivator, MHC2TA, is associated with type 1 diabetes suceptibility. 42nd EASD annual meeting. Copenhagen-Malmoe.

Borkakoti, N. (2004). Matrix metalloprotease inhibitors: design from structure. Biochem Soc Trans **32**: 17-20.

Casiday and Frey, Iron Use and Storage in the Body: Ferritin and Molecular Representations Iron in Biology: Study of the Iron Content in Ferritin, The Iron-Storage Protein. Authors: Rachel Casiday and Regina Frey Department of Chemistry, Washington University St. Louis, MO 63130 For information contact: R. Frey at gfrey@wuchem.wustl.edu.

Chandler, S., K. Miller, et al. (1997). Mateix metalloproteinases, tumor necrosis factor and multiple sclerosis: an overview. J Neuroimmunol **72**: 155-61.

Compston, A. (1997). Genetic epidemiology of multiple sclerosis. J Neurol Neurosurg Psychiatry **62**(6): 553-61.

De Vos, W. & Tarvainen, T. (Chief-editors) & Salminen, R. & Reeder, S. & De Vivo, B. et al. (2006). Geochemical Atlas of Europe. Part 2: Interpretation of Geochemical Maps, Additional Tables, Figures, Maps, and Related Publications. 690 p. ISBN: 951-690-956-6 (printed) & 951-690-960-4 (electronic).

Dore-Duffy, Paula Dr., Dore-Duffy, PhD , Catalanotto, Frank DMD, Donaldson, James O. MD, Ostrom, Karin M. MS, Testa, Marcia A. PhD**,** Zinc in Multiple Sclerosis, in Annals of Neurology, Vol 14, issue 4,pages 450-454 Published Online: Oct 7 2004 9:29AM DOI: 10.1002/ana.410140409.

Haglund, B., Ryckenberg, K., Selinus, O., Dahlqvist, G. (1996). Evidence of a relationship between childhood-onset type 1 diabetes and low groundwater concentration of Zinc. Diabetes Care, Vol 19, No 8, August.

International Council on Mining and Metals, (2007), MERAG: Metals Environmental Riskassessment Guidance, ISBN: 987-0-9553591-2-5, available online and hardcopy from ICMM, www.icmm.com info@icmm.com.

International Multiple Sclerosis Genetics Consortium (IMSGC) (2008). Refining novel associations in multiple sclerosis. Lancet Neurol. 7:567-569.

International Multiple Sclerosis Genetics Consortium (IMSGC) (In Press). The expanding genetic overlap between multiple sclerosis and type 1 diabetes. Genes & Immunity.

Kidd, P. (2003). Th1/Th2 balance: the hypothesis, its limitations, and implications for health and disease. Altern Med Rev **8**: 223-46.

Kockum I, Sanjeevi CB, Eastman S, Landin-Olsson M, Dahlquist G, Lernmark Å. (1995): Population analysis of protection by HLA-DR and DQ genes from insulin-dependent diabetes mellitus in Swedish children with insulin-dependent diabetes and controls. Eur J Immunogenet 22:443-465.

Kurtzke, John F. (2006). An epidemiologic view of etiology of multiple sclerosis, Arch Neurocien (Mex) Vol. 11, No 2:90-99.

Kurtzke, John F. M.D. (1980). Epidemiologic contributions to multiple sclerosis, an overview, Neurology 1980; 30: 61-79.

Landner, Lars and Lindeström, Lennart (1999). Copper in society and the environment, an account of the facts on fluxes, amounts and effects of copper in Sweden, Swedish Environmental Research Group (MFG), ISBN 91-630-7932-1, copies of the book can be ordered from: SCDA, Box 594, SE-721 10 Västerås, Sweden, scda.info@outokumpu.com www.scda.com.

Landtblom, A-M, Riise T, Kurtzke JF. (2005). Further considerations on the distribution of multiple sclerosis in Sweden, Acta Neurol Scand 2005: 111: 238-246 DOI: 10.1111/j.1600-0404.2005.00395.x

Lax, Kaj. (2005). Environmental applications of biogeochemical data from Geological survey of Sweden, Licentiate thesis, Luleå University of Technology, 2005:95, ISSN:1402-1757, ISRN:LTU-LIC--05/95—SE

Lundmark, F., K. Duvefelt, et al. (2007). Variation in interleukin 7 receptor alpha chain (IL7R) influences risk of multiple sclerosis. Nat Genet **39**(9): 1108-13.

Marrie, R. A. (2004). Environmental risk factors in multiple sclerosis aetiology. Lancet Neurol **3**(12): 709-18.

Munger KL, Zhang SM, O'Reilly E, Hernán MA, Olek MJ, Willett WC, & Ascherio A (2004). Vitamin D intake and incidence of multiple sclerosis. Neurology, 62: 60–65.

Nielsen TR, Pedersen M, Rostgaard K, Frisch M and Hjalgrim H, (2007). Correlations between Epstein-Barr virus antibody levels and risk factors for multiple sclerosis in healthy individuals Mult Scler 2007; 13; 420 originally published online Jan 29, 2007; DOI: 10.1177/1352458506071470

Olerup O, Hillert J. (1991). HLA class II-associated genetic susceptibility in multiple sclerosis: A critical evaluation. Tissue Antigens 38:1-15, 1991

Palm, R. , Hallmans, G. (1982). Zinc and copper in multiple sclerosis, Journal of Neurology, Neurosurgery, and psychiatry, volume 45, pages 691-698.

Ponsonby AL, Lucas RM, & van der Mei IA (2005). UVR, vitamin D and three autoimmune diseases — multiple sclerosis, type 1 diabetes, rheumatoid arthritis. Photochem Photobiol, 81(6): 1267–1275.

Ponsonby AL, McMichael A, & van der Mei I (2002). Ultraviolet radiation and autoimmune disease: insights from epidemiological research. Toxicology, 181–182: 71–78.

Riise T, Moen BE, & Kyvik KR (2002). Organic solvents and the risk of multiple sclerosis. Epidemiology, 13: 718–720.

Riise T, Nortvedt MW, & Ascherio A (2003). Smoking is a risk factor for multiple sclerosis. Neurology, 61: 1122–1124.

Robb, Laurence (2008). Introduction to oreforming processes, Blackwell publishing, ISBN 978-0-632-06378-9

Rosati, G. (2001). The prevalence of multiple sclerosis in the world: an update, Neurol Sci (2001) 22:117-139

Salminen, R. (chief-editor) & Batista, M. J. & Bidovec, M. & Demetriades, A. et al.(2005). Geochemical Atlas of Europe. Part 1: Background Information, Methodology and Maps. 525 p. ISBN: 951-690-921-3 (printed) & 951-690-913-2 (electronic).

SOSFS 2003:17, (2003). Försiktighetsmått för dricksvatten, allmänna råd, Socialstyrelsens författningssamling, ISSN 0346-6000 artikelnummer 2003-10-17, in Swedish.

Statens strålskyddsinstitut (2008). Naturligt radioaktiva ämnen, arsenik och andra metaller i dricksvatten från enskilda brunnar, Britt-Marie Ek och Bo Thunholm, SGU; Inger Östergren, Rolf Falk och Lars Mjönes, SSI rapport 2008:15, april 2008, ISSN 0282-4434, www.ssi.se, in Swedish.

Stein, EC, Schiffer, RB, Hall WJ, & Young N (1987). Multiple sclerosis and the workplace: report of an industry-based cluster. Neurology, 37: 1672–1677.

Swanberg, M., Lidman, O. et al. (2005). MHC2TA is associated with differential MHC molecule expression and susceptibility to rheumatoid arthritis, multiple sclerosis and myocardial infarction. Nat Genet **37**(5): 486-94.

Swank RL, Lerstad O, Strom A, et al. (1952). Multiple sclerosis in rural Norway. New Engl J Med 1952;246:721-728.

Todd, J. A., N. M. Walker, et al. (2007). Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. Nat Genet **39**(7): 857-64.

Wagner, S., H. J. Breyholz, et al. (2006). Molecular imaging of matrix metalloproteinases in vivo using molecular inhibitors for SPECT and PET. Curr Med Chem **13**: 2819-2838.

Van Amerongen BM, Dijkstra CD, Lips P, & Polman CH (2004). Multiple sclerosis and vitamin D: an update. Eur J Clin Nutr, 58: 1095–1109.

World Health Organization (2006A). Environmental Health Criteria 236, Principles and methods for assessing autoimmunity associated with exposure to chemicals, ISBN 92 4 157236 1 (NLM classification: WD 305), ISBN 978 92 4 157236 1, ISSN 0250-863X, Publications of the World Health Organization can be obtained from WHO Press, World Health Organization, 20 Avenue Appia, 1211 Geneva 27, Switzerland (tel: +41 22 791 3264; fax: +41 22 791 4857; e-mail: bookorders@who.int).

World Health Organization (2006B). Guidelines for drinking-water quality [electronic resource]: incorporating first addendum. Vol. 1, Recommendations. – 3rd ed. ISBN 92 4 154696 4

World Health Organization, (1998). Environmental Health Criteria 220, Copper, ISBN 92 4 157200 0, ISSN 0250-863X

World Health Organization, (2001). Environmental Health Criteria 221, Zinc, ISBN 92 4 157221 3, ISSN 0250-863X

Zhang SM, Willett WC, Hernán MA, Olek MJ, & Ascherio A (2000). Dietary fat in relation to risk of multiple sclerosis among two large cohorts of women. Am J Epidemiol, 152: 1056–1064.

8.2 References on multivariate analysis

Beebe K.R. and Kowalski R.B. (1987). "An Introduction to Multivariate Calibration and Analysis", *Anal. Chem.* **59**, 1007A-1017A. (PLS)

Geladi P. and Kowalski B.R.(1986). "Partial Least-Squares Regression : A Tutorial", *Anal. Chim. Acta* **185**, 1-17. (PLS)

Martens, H. and Naes, T. (1989). *Multivariate Calibration*, *Multivariate Calibration*. 700p., John Wiley & Sons, N.Y. (PLS/PCA)

Trygg J, Wold S. Orthogonal projections to latent structures, O-PLS.(2002). J. Chemometrics, 16: 119-128. [http://www3.interscience.wiley.com/cgibin/ abstract/89015377/]

Wold S., Esbensen K. and Geladi P.(1987). "Principal Component Analysis", *Chemometr. Intellig. Lab. Syst.* **2**, 37-52.

8.3 Internetlinks

Anon., Multiple Sclerosis Resource Centre BBD Supplements Explained, http://www.msrc.co.uk/index.cfm?fuseaction=show&pageid=987, [2008-07-18]

Peter Betts Good, Multiple Sclerosis Studies, http://www.goldwarp.com/MultipleSclerosisStudies/index.html, Bend, OR, USA, [2008-07-29]

Rachel Casiday and Regina Frey, The Crucial Role of Iron in the Body, http://www.chemistry.wustl.edu/~courses/genchem/Tutorials/Ferritin/IronBody.htm, St. Louis, MO, USA, [2008-07-29]

8.4 Software

ESRI ArcInfo Workstation 9.2

Multivariate program SIMCA-P+12.0 from Umetrics AB, Umeå

The MathWorks Inc. MatLab 7.4