ORIGINAL PAPER

Spatial correlation between the prevalence of transmissible spongiform diseases and British soil geochemistry

C. E. Imrie · A. Korre · G. Munoz-Melendez

Abstract Transmissible spongiform encephalopathies (TSEs) are a group of fatal neurological conditions affecting a number of mammals, including sheep and goats (scrapie), cows (BSE), and humans (Creutzfeldt-Jakob disease). The diseases are widely believed to be caused by the misfolding of the normal prion protein to a pathological isoform, which is thought to act as an infectious agent. Outbreaks of the disease are commonly attributed to contaminated feed and genetic susceptibility. However, the implication of copper and manganese in the pathology of the disease, and its apparent geographical clustering, have prompted suggestions of a link with trace elements in the environment. Nevertheless, studies of soils at regional scales have failed to provide evidence of an environmental risk factor. This study uses geostatistical techniques to investigate the correlations between the distribution of TSE prevalence and soil geochemical variables across the UK according to different spatial scales. A similar spatial pattern in scrapie and BSE occurrence is identified, which may be linked with increasing pH and total organic carbon, and decreasing iodine concentration. However, the pattern also resembles that of the density of dairy farming. Nevertheless, despite the low spatial resolution of the TSE data available for this study, the fact that significant correlations are detected indicates there is a possibility of a link between soil geochemistry, scrapie, and BSE. It is suggested that further investigations of the prevalence of TSE and environmental exposure to trace metals should take into account the factors affecting their bioavailability.

Keywords Transmissible spongiform diseases · Soil geochemistry · Scrapie · BSE · Geostatistics · Linear model of coregionalisation · Prion

Introduction

Transmissible spongiform encephalopathies (TSEs) are a group of fatal neurological conditions affecting a number of mammals. They include: scrapie, which affects sheep and goats; bovine spongiform encephalopathy (BSE), which affects cows; chronic wasting disease (CWD), which mainly occurs in deer and elk; and Creutzfeldt-Jakob disease (CJD), which affects humans. The diseases are widely believed to be caused by the misfolding of the normal prion protein (PrP^c) to a pathological isoform (PrP^{Sc}), which is thought to act as an infectious agent (Prusiner 1982). While there is evidence that PrP^c acts as a superoxide dismutase in normal tissue (Brown et al. 1999), there is no evidence of a means of transmission; a number of researchers are therefore contesting this belief with

C. E. Imrie (✉) · A. Korre · G. Munoz-Melendez
Department of Earth Science and Engineering, Imperial College London, London SW7 2AZ, UK
e-mail: c.imrie@imperial.ac.uk

alternative theories. Such propositions include the involvement of a bacterial agent (e.g. Broxmeyer 2004; Ebringer et al. 2005; Bastian 2005), while Manuelidis (2007) is a strong proponent of the theory that TSEs are caused by a unique class of "slow" viruses. Others (e.g. Purdey 1998; Gordon et al. 1998) have suggested that the agricultural use of organophosphate pesticides is implicated. Georgsson et al. (2006) collected evidence that suggested the infectious agent may persist in the environment for more than a decade. A link between TSEs and trace metals in soils (e.g. Purdey 2000) has also been suggested, and this has prompted the investigation reported in this paper.

Interest in environmental geochemistry as a risk factor for TSEs arose partly from indications of geographical clustering. For example, scrapie in Iceland is known to occur on certain farms or valleys and not others (Sigurdarson 2000). Clusters of TSEs have also been reported for variant CJD in Leicestershire, UK, BSE in Switzerland (Doherr et al. 2002). and CWD in the Rocky Mountains. Tongue et al. (2006) identified three statistically significant clusters with a relatively high risk of scrapie for sheep flocks in the central south, north-east, and far north of England, and low-risk clusters were located in north-west Scotland, north-west Wales and south Yorkshire.

While clustering is often attributed to genetic factors, exposure to contaminated feedstuffs, and differential reporting, a link with trace metals is supported by evidence that on replacement of copper with manganese as the metal ligand of the prion protein, the protein's structure is altered (Brown et al. 2000). Moreover, Wong et al. (2001) reported a tenfold increase in manganese and a 50% reduction in copper in the brains of victims of CJD. More recently, Hesketh et al. (2007) found that elevated levels of manganese occur in the blood and central nervous system of sheep and cows prior to the onset of clinical signs of scrapie and BSE.

Attempts to link TSEs with trace elements in the environment were initiated by Purdey (2000), who sampled soils from TSE-affected areas across the world. While localised studies indicated lower levels of copper and higher concentrations of manganese on scrapie-affected farms in comparison with scrapie-free farms, Chihota et al. (2004) found no evidence for such a relationship occurring on a regional scale across the British Isles. An investigation into the possible connection between scrapie in Iceland and the levels of selenium in hay also proved inconclusive (Jóhannesson et al. 2004). However, preliminary results from a study of trace metals in soil samples from scrapie-affected areas in Iceland indicated high amounts of bioavailable manganese and low amounts of soluble and free copper (Ragnarsdottir and Hawkins 2006). On investigating potential biogeochemical risk factors for a CWD cluster in Wisconsin, McBride (2007) found high levels of sulfur in a willow species often used for forage; this may have contributed towards copper deficiency.

Despite the interest in the involvement of environmental factors in the epidemiology of TSEs, no geographical correlation between the occurrence of the different TSEs has ever been established. The individual diseases are commonly associated with specific traits which may serve to mask any epidemiological correlation. For example, scrapie is known to have a higher prevalence among specific genotypes (in particular the ARR allele), and for this reason the National Scrapie Plan, designed to reduce and eventually eliminate scrapie by selective breeding (focussing on the VRQ allele), was introduced in the UK in 2001 (DEFRA 2006a). Variant CJD has also been found to occur exclusively in victims who share the MM codon 129 genotype, and this has been found in around 64% of cases of sporadic CJD to date (National CJD Surveillance Unit 2007).

Although a slight genetic susceptibility of cattle to BSE has recently been demonstrated (Juling et al. 2006; Haase et al. 2007), it is assumed to be associated primarily with dairy farming, because of exposure of non-grazing ruminants to meat and bone meal (MBM). Approximately 81% of BSE cases to date have occurred in dairy cattle (DEFRA 2006b). After the implementation of the MBM feed ban in 1988, ongoing occurrences of BSE were linked with cross-contamination with feed intended for non-ruminants (e.g. Stevenson et al. 2005). However, the cases occurring after the reinforcement and extension of the ban to include all offal, in 1996, have yet to be explained. Because of the timing of its emergence, new-variant CJD has been linked with the consumption of contaminated beef. Nevertheless, it remains possible that a common environmental factor presents an additional risk.

In 2003, a European Commission-funded project (FATEPriDE, QLK4-CT2002-02723) investigating whether environmental factors influenced the occurrence of TSEs commenced. The main objective of the project focussed on the possible role of trace elements and, in particular, the balance between manganese and copper (Brown 2007), and involved laboratory experimentation, field studies, and computer modelling. This included an analysis of the geographical distribution of manganese, copper, and other trace elements across Europe and a comparison of the resulting geochemical maps with those indicating high TSE incidence rates. This study used geochemical data recently collected by the FORum of European Geological Surveys (Salminen et al. 2005) from topsoils and other media across 26 participating European countries, at a resolution of around 72×72 km^2. Part of this investigation focussed on the British Isles, because of the availability of regional incidence data. This paper describes the research undertaken in this study and discusses inferences from the results.

Materials and methods

Data

The data on scrapie and BSE incidence in the 65 administrative counties of Scotland, England, and Wales were obtained from DEFRA. Data on confirmed cases of scrapie in sheep and goats were available through the Scrapie Notifications Database covering the period from January 1998 to September 2002. The total numbers of BSE cases per county confirmed by passive surveillance were also available, covering the emergence of the disease in November 1986 up to the 6 January 2006. Unfortunately these data do not reveal the changing geographical distribution of BSE over time detected by Stevenson et al. (2005) and attributed to the feed-ban abatement measures. However, they do preserve the prevailing north–south trend. The scrapie and BSE data were converted to annual incidence rates using the agricultural census data from each county for the year 2000. The census data were also used to calculate the percentage of the cattle in each county that were used for dairy production.

Numbers of deaths from definite and probable sporadic (sCJD) and new-variant (vCJD) CJD per county from 1 January 1990 to 31 December 2004 were obtained from the National CJD Surveillance Unit (2005). These were converted to incidence rates per county using national census data collected in 2001.

The incidence rates calculated for the four TSEs in each county were compiled in a GIS database and mapped, along with the percentage of dairy cattle (Fig. 1a–e). The maps were plotted according to a decile classification.

Soil geochemistry data were obtained from the FOREGS database (Salminen et al. 2005). This dataset was compiled following a standardised sampling and analysis campaign across Europe. As such, its spatial resolution is somewhat coarse (roughly one sample per 5,000 km^2), with a total of 845 topsoil samples collected throughout the 26 participating countries. The survey was aimed at establishing baseline levels of a wide range of major and minor elements (except selenium), pH, and organic content of soils, stream waters, stream sediments, and humus. Topsoil samples, which were taken from a depth of 0 to 25 cm, have been used in this study. For more information regarding the sampling and analysis techniques consult Salminen et al. (1998).

Care was taken during the FOREGS sampling campaign to avoid selecting locations that may have been subject to anthropogenic contamination, and preference was given to forested and unused lands, pastures, meadows and agricultural land (Salminen et al. 1998). Multivariate geostatistical research conducted on the entire FOREGS topsoil data has demonstrated that they are representative of regional background soil geochemistry, as governed by parent material geology and soil-formation processes, over local, intermediate, and long-range spatial scales (Imrie et al. 2008).

The trace elements included in this study were copper (Cu), manganese (Mn), zinc (Zn), molybdenum (Mo), iodine (I), iron (Fe), and sulfur (S). These elements were selected on the basis of a comprehensive literature review and the concurrent research and findings in the FATEPriDE project. The concentrations of the elements were analysed using exhaustive extraction methods (aqua regia digestion followed by measurement with inductively coupled plasma (ICP) mass spectrometry for Cu, Mn, Zn, Fe, and S, and
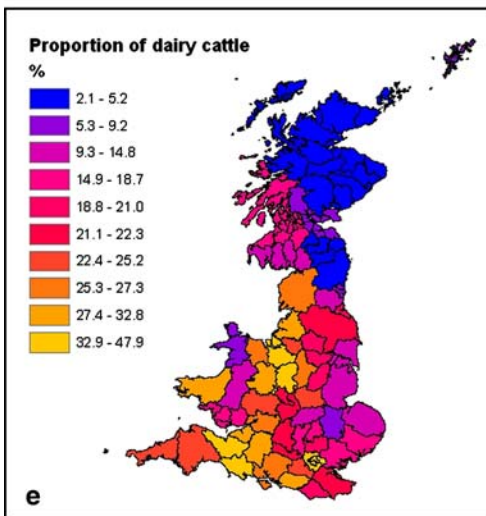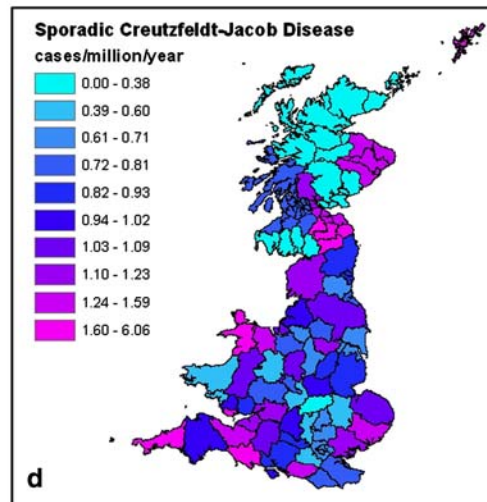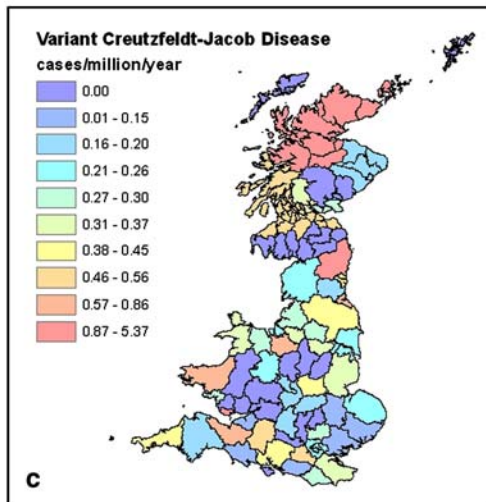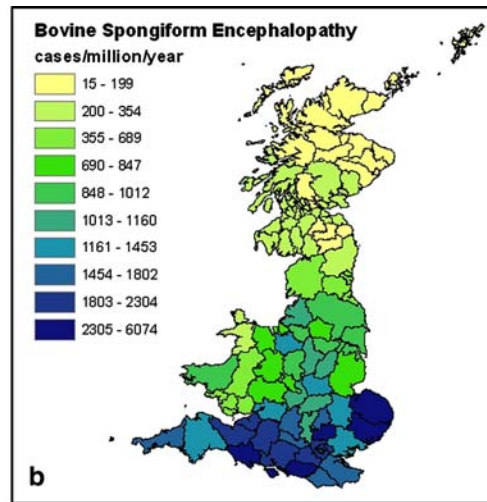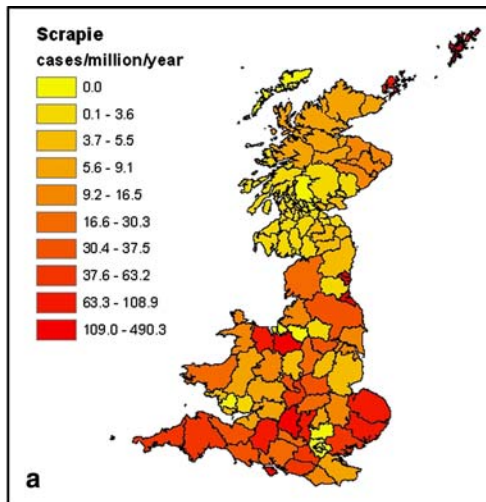
◄**Fig. 1** Distribution of incidence rates of scrapie (**a**), bovine spongiform encephalopathy (**b**), variant Creutzfeldt-Jacob disease (**c**), and sporadic Creutzfeldt-Jacob disease (**d**), and the percentage of cattle used for dairy production (**e**), across the British Isles

mixed acid digestion followed by ICP atomic emission spectrometry for I and Mo), and so can be assumed to represent their total concentration. Additional variables included were total organic carbon (TOC), pH, and the clay fraction (CF: percentage of soil particles <0.002 mm in diameter). These variables are known to affect the "bioavailability" of trace elements in the soil and hence their uptake by vegetation. For more information regarding the analysis of the variables see Salminen et al. (2005).

Methods

As noted above, various independent factors are associated with the incidence of scrapie, BSE, and vCJD. Such factors are likely to mask correlations between the TSEs unless they can be filtered out. If it may be assumed that the geographical distributions of genetic and feed-based causative factors are unrelated to the geographical distribution of trace element levels across the British Isles, then a multivariate spatial analysis of the data can be used to isolate their individual contributions.

Multivariate geostatistics enables the spatial analysis of two or more collocated variables, whereby the cross-correlations between pairs of variables can be estimated according to the various spatial scales identified in the dataset. For example, it would be possible for two geochemical variables to be positively correlated in the micro-scale spatial range because of a common sampling or analytical bias, and negatively correlated over longer ranges because of differing geochemical attributes (for example affinity for organic carbon, or relative concentration in a certain parent material geology).

The spatial autocorrelation of a variable is represented graphically by a variogram, as described in more detail in the Appendix. A variogram model may consist of one or more structures with different spatial scales (ranges), including a nugget effect which represents variation occurring at a range shorter than the sampling density. Variograms are used to assign weights to neighbouring samples in order to provide unbiased estimates of the mean value of variable at a

location of interest. This estimation, known as kriging, may also be used to find mean values of a variable over a regular grid (block kriging) or over irregular shapes (polygon kriging).

A linear model of coregionalisation (LMC) is a set of variograms and cross-variograms calculated for two or more collocated variables. For each spatial scale identified, a covariance matrix may be generated, from which correlation coefficients between variables can be estimated. The methods are described in more detail in the Appendix.

The methodology followed in this study is briefly outlined below.

1. *Collocation of datasets*: To create an LMC the data need to be collocated. Because the geochemical data are derived from point samples and the TSE incidence data are available at county level, the mean values of geochemical data must be estimated for each of the administrative divisions. This is done using polygon kriging.
2. *Identification of spatial structures*: Variograms are calculated and modelled for each of the variables, using the centroid of each county to define the locations of the data. The number and ranges of the common spatial structures are identified.
3. *Linear model of coregionalisation*: An LMC is fitted to the set of variograms and cross-variograms, according to the number and ranges of the spatial structures identified. The model is fitted using a weighted least-squares approach.
4. *Calculation of correlation coefficients*: A covariance matrix is obtained for each spatial structure, and used to calculate structural correlation coefficients between each of the variables.

Spatial analysis and results

The FOREGS geochemical database includes soil samples taken from 56 sites across the British Isles. Although this is less than the number of administrative divisions, research based on the FOREGS data has shown that there is significant spatial continuity in topsoil geochemistry across the whole of Europe (Imrie et al., 2008). Variograms were calculated and modelled for each of the selected variables. Polygon

kriging was then used to estimate the mean concentrations of the variables in each of the counties. Example maps of the distribution of pH, TOC, CF, Cu, I, and Mn, plotted according to a decile classification, are provided in Fig. 2. Unfortunately it was necessary to omit the Shetland, Orkney, and Western Isles from this analysis because of the lack of soil samples taken from these areas and the high uncertainty in the kriging estimates.

Inspection of the maps in Fig. 1 shows that there is a strong north–south trend for BSE, and similar but weaker trends for scrapie and the percentage of cattle used for dairy (PCD). The distributions of the variables in Fig. 2 show similarities between some of the soil constituents, such as I and total organic content. Cu, Mn, pH, and the clay fraction show a general north–south trend, although there are some longitudinal variations.

Variant and sporadic CJD are more randomly distributed in comparison. The global correlation coefficients were calculated between the TSEs, the PCD and the geochemical levels for each county, and these are provided in Table 1. The variables were first log transformed where necessary to reduce the skewness of the data.

Table 1 reveals significant correlations between BSE and soil pH, Cu, Zn, and PCD. Weaker correlations are observed for scrapie and I, TOC, and PCD, and there is a weak correlation between scrapie and BSE (0.371). There is little correlation between these TSEs and sporadic and variant CJD.

A full set of variograms and cross-variograms was calculated and a linear model of coregionalisation was fitted. Three spatial structures were identified: a nugget effect, a short-range structure with a range of 130 km, and a long-range structure (range >130 km). The structural components were fitted using spherical variogram functions. Examples of the variograms of scrapie and I, and their covariogram, are provided in Fig. 3. The LMC was then used to compute a variance–covariance matrix, from which correlation coefficients corresponding to each spatial scale were calculated. The results are given in Table 2.

It is also important to consider the amount of variance that is contained within each spatial range component. For example, most of the variance in scrapie incidence is contained within the nugget component (62%) while 17% and 21% are contained

Fig. 2 Distribution of (**a**) pH, (**b**) total organic carbon, (**c**) clay ▶ fraction, (**d**) copper, (**e**) iodine, and (**f**) manganese in topsoils across the British Isles

in the short and long-range components, respectively. In contrast, the BSE variance is mostly contained in the long-range component (82%), with 5% and 13% in the nugget and short-range components respectively. Correlation coefficients calculated for variables in which the overall variance is low may be considered unreliable. The percentages of variances contained in the respective structural components are given in brackets in the column headings in Table 2.

The nugget effect on a variogram contains variance that may be due to measurement error and variation on a scale smaller than that which can be captured by the spatial resolution of the data. Variables which have a particularly high nugget component are scrapie incidence rate, Mo, and vCJD, with sCJD, S and TOC having lower but still significant nugget components. Among these variables, scrapie incidence is negatively correlated with TOC and S.

Iodine is most variable within the short range (<130 km) spatial component; Mo, S, TOC, pH, PCD, and sCJD vary to a lesser extent. Although scrapie and BSE incidence vary by 17% and 13% of their overall variance, respectively, there is a strong positive correlation between them (0.705). There is also a correlation between sCJD and vCJD over this spatial range. The clay fraction, Cu, I, Mn, Fe, and Mo are all negatively correlated with both scrapie and BSE, while pH is positively correlated. BSE has a moderate correlation with the percentage of cattle used for dairy production.

Most of the variance in BSE incidence, CF, Cu, Fe, Mn, Zn, pH, and PCD is contained in the long-range component (>130 km). There is a strong correlation again between BSE and scrapie incidence rates (0.812). Both TSEs are positively correlated with CF, Cu, Fe, Zn, pH, and PCD, and negatively correlated with I and TOC. While BSE was positively correlated with vCJD over the short spatial range, the variables are negatively correlated over the long spatial range.

Discussion

The results obtained from the linear model of coregionalisation have indicated a regional correlation between BSE and scrapie incidence. Relationships
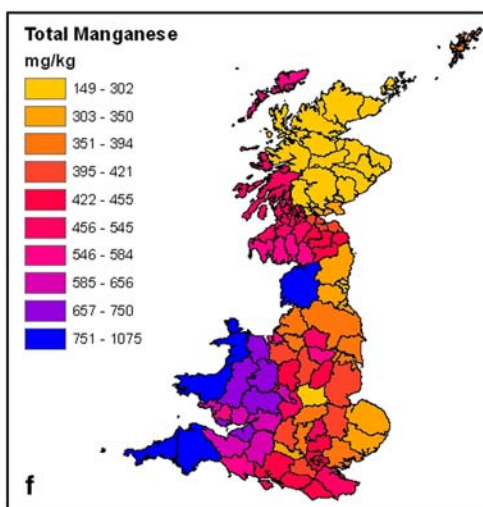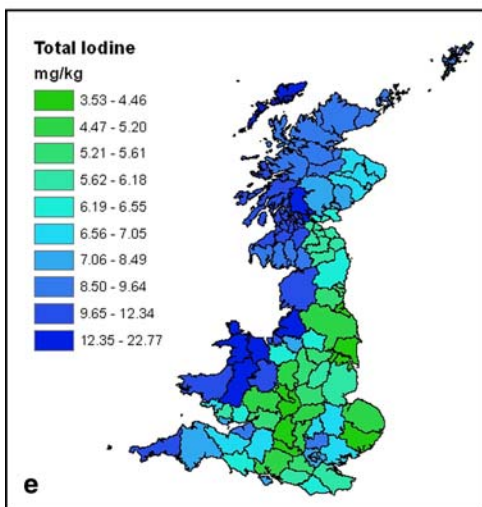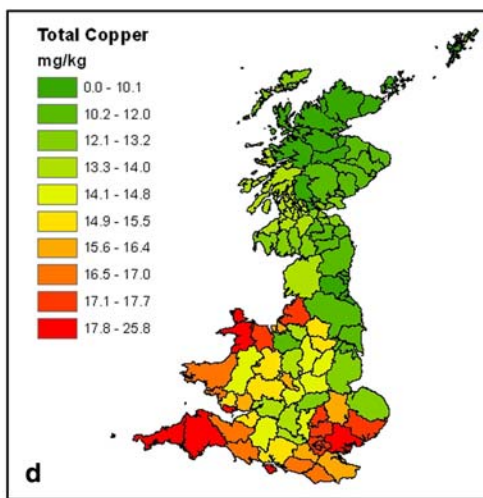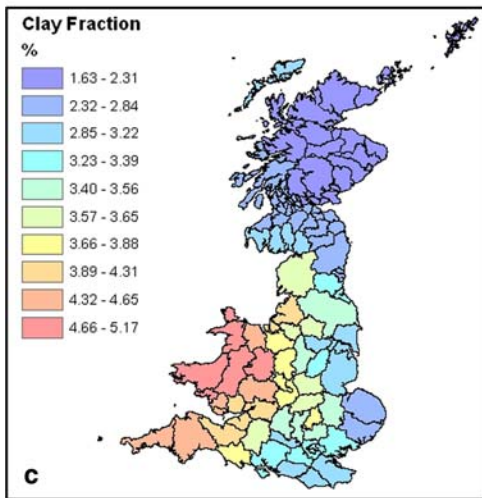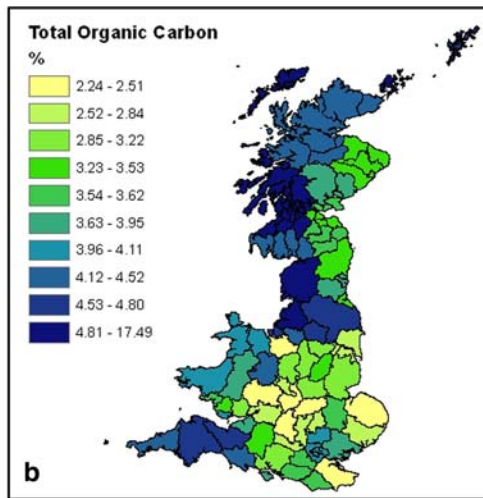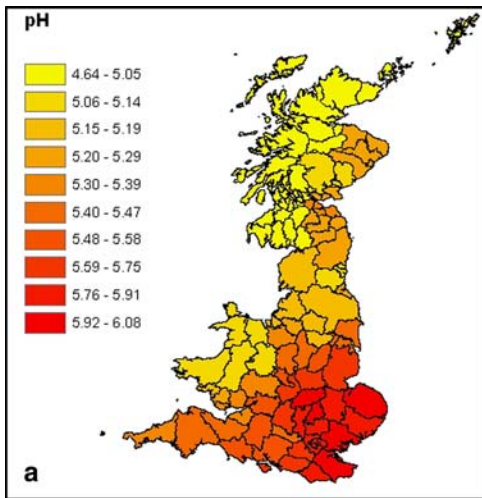
Table 1 Global correlation coefficients calculated between the TSEs (sporadic Creutzfeldt-Jacob disease (sCJD); variant Creutzfeldt-Jacob disease (vCJD); scrapie; and bovine spongiform encephalopathy (BSE)), the percentage of cattle used for dairy (PCD), and the geochemical variables (clay fraction (CF), Cu, Fe, I, Mn, Mo, S, total organic carbon (TOC), Zn, and pH)

	CF	Cu	Fe	I	Mn	Mo	S	TOC	Zn	pH	PCD	sCJD	vCJD	Scrapie	BSE
sCJD	0.170	0.279	0.048	0.184	0.241	0.098	0.216	0.140	0.080	−0.050	0.000	1.000			
vCJD	−0.103	−0.179	−0.145	0.172	−0.057	0.105	0.062	0.200	−0.325	−0.299	−0.123	0.025	1.000		
Scrapie	0.042	0.053	0.019	−0.336	−0.042	−0.173	−0.179	−0.265	0.146	0.258	0.275	0.117	0.054	1.000	
BSE	0.294	**0.507**	0.197	−0.311	0.187	−0.298	0.071	−0.194	**0.596**	**0.726**	**0.617**	0.073	−0.180	0.371	1.000

The highest correlation coefficients have been highlighted in bold type
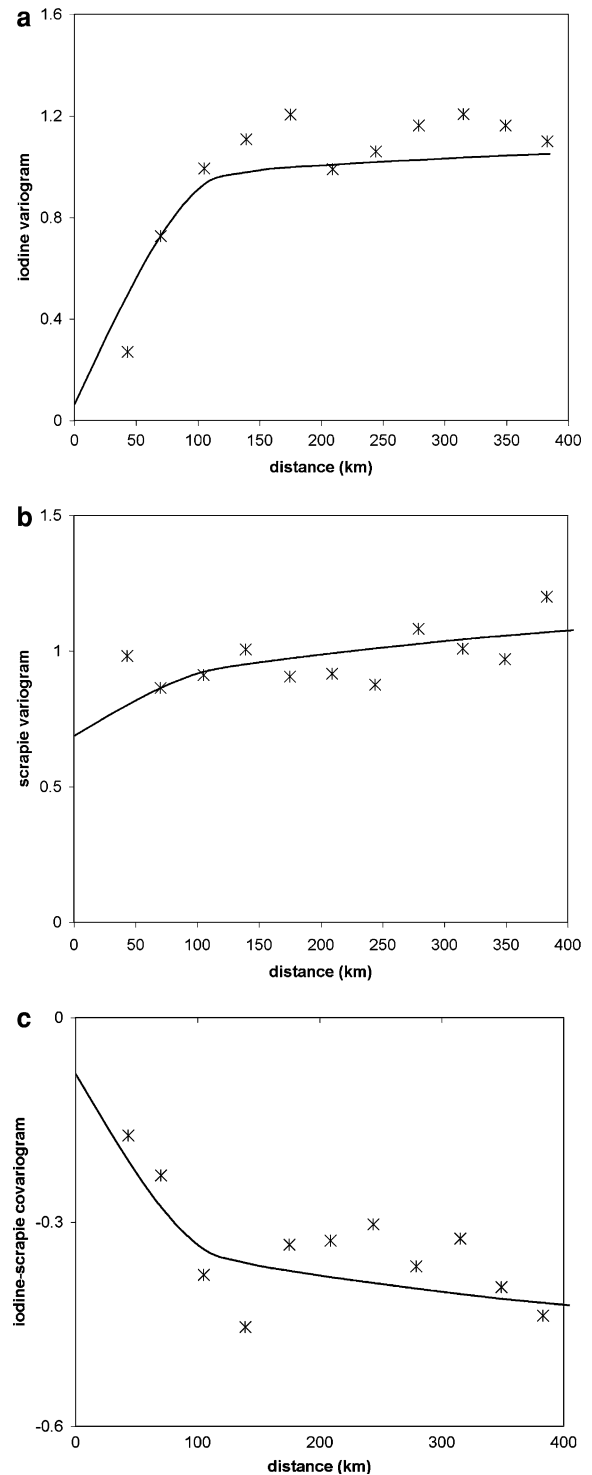
Fig. 3 Variograms of topsoil iodine (**a**), scrapie incidence (**b**) and their covariogram (**c**). The variables were normalised and standardised prior to calculation of the variograms

Table 2 Structural correlation coefficients calculated between the TSEs (sporadic Creutzfeldt-Jacob disease (sCJD); variant Creutzfeldt-Jacob disease (vCJD); scrapie; and bovine spongiform encephalopathy (BSE)), the percentage of cattle used for dairy (PCD), and the geochemical variables (clay fraction (CF), Cu, Fe, I, Mn, Mo, S, total organic carbon (TOC), Zn and pH) for each spatial range

Nugget	CF (2)	Cu (10)	Fe (16)	I (6)	Mn (4)	Mo (61)	S (42)	TOC (26)	Zn (13)	pH (3)	PCD (4)	sCJD (29)	vCJD (67)	Scrapie (62)	BSE (5)
sCJD	-0.576	0.151	-0.122	-0.029	0.003	0.085	0.023	-0.001	0.156	-0.121	0.188	1.000			
vCJD	-0.114	0.124	0.080	-0.161	0.005	0.042	-0.057	-0.237	0.150	0.111	-0.170	-0.048	1.000		
Scrapie	-0.038	-0.353	0.070	-0.388	-0.273	0.067	-0.410	**-0.402**	-0.260	-0.532	0.125	0.077	0.173	1.000	
BSE	-0.033	-0.372	-0.062	0.441	-0.446	-0.236	0.004	0.001	-0.272	-0.470	0.573	-0.134	-0.170	-0.125	1.000

<130 km	CF (20)	Cu (12)	Fe (17)	I (82)	Mn (15)	Mo (30)	S (35)	TOC (40)	Zn (10)	pH (21)	PCD (30)	sCJD (25)	vCJD (13)	Scrapie (17)	BSE (13)
sCJD	0.416	0.544	-0.127	0.517	0.657	0.377	**0.724**	**0.794**	-0.507	-0.704	0.078	1.000			
vCJD	0.120	0.381	-0.573	0.489	0.181	0.017	**0.844**	**0.864**	-0.841	-0.403	**0.697**	**0.580**	1.000		
Scrapie	**-0.726**	-0.426	-0.445	**-0.596**	-0.468	**-0.756**	-0.141	-0.211	-0.059	0.453	-0.131	0.184	0.008	1.000	
BSE	**-0.729**	-0.573	**-0.857**	**-0.552**	-0.519	**-0.688**	-0.047	-0.047	-0.562	0.385	0.432	0.052	0.362	**0.705**	1.000

>130 km	CF (78)	Cu (78)	Fe (67)	I (12)	Mn (81)	Mo (9)	S (23)	TOC (34)	Zn (77)	pH (76)	PCD (66)	sCJD (46)	vCJD (20)	Scrapie (21)	BSE (82)
sCJD	0.028	0.171	0.212	-0.449	0.092	-0.538	0.079	-0.326	0.175	0.404	0.044	1.000			
vCJD	-0.328	-0.518	-0.193	0.329	-0.279	0.274	0.031	0.532	-0.574	-0.584	-0.487	0.077	1.000		
Scrapie	**0.778**	**0.576**	0.358	-0.656	0.315	-0.057	-0.059	**-0.742**	**0.824**	**0.722**	**0.735**	0.057	**-0.574**	1.000	
BSE	**0.608**	**0.866**	0.611	-0.488	0.410	-0.415	-0.158	**-0.782**	**0.881**	**0.892**	**0.762**	0.095	**-0.595**	**0.812**	1.000

The numbers in parentheses refer to the percentage of variance contained in each spatial structure, and the highest correlation coefficients have been highlighted in bold type

between these TSEs and sporadic and variant CJD, however, are less clear.

The lack of significant correlation between scrapie and BSE in the nugget component may be due to local variations in genetic susceptibility (breeds of sheep) and exposure to contaminated feedstuffs (cattle). The strong correlations observed over the short and long spatial ranges may have several causes, including a common geochemical factor. It is, however, clear that the distribution of the percentage of dairy cattle shares common features with the soil geochemistry of the British Isles. This is somewhat unfortunate, as it is difficult to distinguish any risk factor from soil geochemistry from that of the exposure of cattle to MBM with the data used for this study. However, it seems unlikely that the PCD should have an impact on the occurrence of scrapie. A possible explanation for the north–south trend in scrapie may be related to flock management practices (such as hill farming in the north) and resulting differences in disease diagnosis and reporting (e.g. Tongue et al. 2006).

If incidence rates of scrapie and BSE are influenced by soil geochemistry, the most important variables would appear to be I, TOC, and pH. While Mn has a moderate positive correlation, Cu (and Zn) is also positively correlated in the long-range component, which would appear to refute the notion that Cu deficiency is a factor. However, it is important to note that the "bioavailable" concentrations of Cu and other trace elements in soils are dependent on factors such as pH, clay content, and TOC (e.g. Kabata-Pendias 2001), while the availability of Mn is also strongly influenced by redox conditions. In general, the availability of trace elements for uptake by vegetation decreases with increasing pH (except Mo) and clay content. Many trace elements in soils, particularly I, have a strong affinity for organic matter. Therefore, it is possible that the correlations with pH, clay fraction, and TOC are related to trace element deficiencies.

Iodine is consistently negatively correlated with scrapie and BSE, as the nugget component of BSE is too small to be of any significance. Iodine deficiency has been linked with that of Se with regard to the levels of glutathione peroxidase produced by the thyroid. Glutathione peroxidase acts as an antioxidant, and has been implicated in the pathology of scrapie (e.g. Zeng et al. 2003; Jóhannesson et al. 2004). Iodine deficiency

is less likely to be a contributory factor in dairy than in beef suckler cases, because dairy cattle routinely receive supplements of I and other micronutrients in their feed.

It is difficult to detect any relationships between CJD and the other variables. However, although scrapie/BSE and sCJD are uncorrelated over the long range, sCJD has similar, if weaker, correlations with topsoil I, pH, and TOC. BSE and vCJD are negatively correlated except for the short-range component where little of their variance is contained.

Although the results have indicated a long-range correlation between scrapie and BSE in the British Isles, it must be emphasised that the study has been performed using limited datasets with coarse spatial resolution. As such, the uncertainties are high and the estimated correlations must be regarded with caution. Furthermore, if the Shetland data had been included, this may have reduced the effect of the north–south trend in the scrapie data. Nevertheless, the results indicate that there is a possibility of a link between scrapie, BSE, and soil geochemistry, and it may be recommended that future investigations into environmental exposure to trace metals take into account factors which determine their mobility, such as total organic carbon, pH, and clay fraction.

Appendix

A.1 Geostatistical estimation

In brief, geostatistics is a set of tools which exploits the spatial correlation structure in a dataset to make unbiased estimates of one or more variables at unsampled locations. In its simplest form, the correlation structure is assessed using the variogram $\gamma(\mathbf{h})$, calculated by pooling pairs of data $\mathbf{x}_i$ into specified intervals of separation distance $\mathbf{h}$. The differences between the values of the data in each pair are

calculated, and for each distance category the mean of the squared subtractions is obtained:

$$\gamma(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{i=1}^{N(\mathbf{h})} [z(\mathbf{x}_i) - z(\mathbf{x}_i + \mathbf{h})]^2 \qquad (1)$$

where $N(\mathbf{h})$ is the number of pairs within a separation distance interval.

A model is then fitted to the experimental variogram to obtain a smooth function with which to calculate appropriate weights during the subsequent estimation procedure. Valid models which are commonly fitted to the experimental variogram include spherical, Gaussian, and exponential functions. These are characterised by a *sill*, which represents the covariance accounted for by the model, and a *range*, which signifies the extent of spatial correlation. The models can be nested, whereby correlation structures at different scales are represented. For example, a typical variogram for geochemical data may comprise the following components:

- a nugget effect which accounts for the covariance in the data arising from measurement and sampling errors, and variation that is on a scale that is too small to be detected by the sampling density;
- a short-scale correlation structure arising from local variations, because of factors such as land use and parent material geology;
- a longer-scale correlation structure or trend arising from, for example, larger scale geological features and climate.

In its simplest form, geostatistical estimation is undertaken using a procedure called ordinary kriging. A neighbourhood is first specified which determines the minimum and maximum number of data points to be used for the estimations, how they should be distributed around the location, and what is the maximum separation distance. The choice of these parameters is generally determined by the amount and density of the data and the characteristics of the variogram. The kriging estimate Z^* is defined as a linear combination of the neighbouring data Z_α:

$$Z^* = \sum_\alpha \lambda^\alpha Z_\alpha \qquad (2)$$

where λ_α are weights derived from the variogram model by minimising the expected estimation error. A comprehensive description of geostatistical estimation may be found in Isaaks and Srivastava (1989) and Goovaerts (1997).

While point kriging is used to estimate values at a point location, block kriging is used to compute an average value over a volume. Polygon kriging can be used to estimate average values over an irregular shape.

A.2 Linear model of coregionalisation

A useful tool for analysing regional geochemistry is the linear model of coregionalisation (LMC). This is essentially a set of variograms and cross-variograms calculated for two or more collocated variables. The cross-variograms are calculated by comparing pairs of different variables with increasing distance. It is therefore possible to detect any spatial scales over which the variables are correlated because of a common underlying factor.

The theory underpinning multivariate geostatistical analysis has been published in a number of texts (e.g. Goovaerts 1997; Wackernagel 1998). A brief outline is provided below.

Cross-variograms between two variables z_α and z_β may be calculated as follows:

$$\gamma_{\alpha\beta}(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{i=1}^{N(\mathbf{h})} [(z_\alpha(\mathbf{x}_i) - z_\alpha(\mathbf{x}_i + \mathbf{h})) \cdot (z_\beta(\mathbf{x}_i) - z_\beta(\mathbf{x}_i + \mathbf{h}))] \qquad (3)$$

As mentioned above, variogram models may be nested, so that they represent a number of processes acting at different spatial scales. Such a covariogram is a linear combination of N_s component functions $g^u(\mathbf{h})$:

$$\gamma_{\alpha\beta}(\mathbf{h}) = \sum_{u=1}^{N_s} \gamma_{\alpha\beta}^u(\mathbf{h}) = \sum_{u=1}^{N_s} b_{\alpha\beta}^u g^u(\mathbf{h}) \qquad (4)$$

where the $b_{\alpha\beta}^u$ are coefficients which represent the importance of each spatial scale u to the relationships between the variables. This LMC can be expressed in matrix terms:

$$\Gamma(\mathbf{h}) = \sum_{u=1}^{N_s} \mathbf{B}^u g^u(\mathbf{h}) \qquad (5)$$

where $\Gamma(\mathbf{h})$ is the $p \times p$ variogram matrix and $\mathbf{B}^u$ is a positive semi-definite matrix of the coefficients $b_{\alpha\beta}^u$. A measure of the correlation between the variables Z_α and Z_β at the spatial scale u is given by:

$$\rho_{\alpha\beta}^u = \frac{b_{\alpha\beta}^u}{\sqrt{b_{\alpha\alpha}^u b_{\beta\beta}^u}} \qquad (6)$$

The structural correlation coefficients $\rho^u_{\alpha\beta}$ differ from the traditional product–moment correlation coefficients in that they focus on specific spatial scales, filtering out the processes operating over different distances. It may also be noted that they are derived within a probabilistic framework and not directly from the data pairs.

References

Bastian, F. O. (2005). Spiroplasma as a candidate agent for the transmissible spongiform encephalopathies. *Journal of Neuropathology and Experimental Neurology, 64*, 833–838.

Brown, D. R., Wong, B. S., Hafiz, F., Clive, C., Haswell, S. J., & Jones, I. M. (1999). Normal prion protein has an activity field like that of superoxide dismutase. *Biochemical Journal, 344*, 1–5.

Brown, D. R., Hafiz, F., Glasssmith, L. L., Wong, B. S., Jones, I. M., Clive, C. C., et al. (2000). Consequences of manganese replacement of copper for prion protein function and proteinase resistance. *EMBO Journal, 19*, 1180–1186.

Brown, D. R. (2007). FATEPRIDE. SEAC 97/4. Spongiform Encephalopathy Advisory Committee. London. http://www.seac.gov.uk/agenda/agen100507.htm. (Data of last access February 1, 2008.)

Broxmeyer, L. (2004). Is mad cow disease caused by a bacteria? *Medical Hypotheses, 63*, 731–739.

Chihota, C. M., Gravenor, M. B., & Baylis, M. (2004). Investigation of trace elements in soil as risk factors in the epidemiology of scrapie. *Veterinary Record, 154*, 809–813.

DEFRA (2006a). *National scrapie plan for Great Britain: scheme brochure*. London: Department for Environment, Food and Rural Affairs on behalf of the GB Agriculture Departments.

DEFRA (2006b). BSE: statistics-general statistics for GB as at 3 January 2006. www.defra.gov.uk/animalh/bse/statistics/bse/general.html. Accessed 13 February 2006.

Doherr, M. G., Hett, A. R., Rüfenacht, J., Zurbriggen, A., & Heim, D. (2002). Geographical clustering of cases of bovine spongiform encephalopathy (BSE) born in Switzerland after the feed ban. *Veterinary Record, 151*, 467–472.

Ebringer, A., Rashid, T., & Wilson, C. (2005). Bovine spongiform encephalopathy, multiple sclerosis, and Creutzfeld-Jakob disease are probably autoimmune diseases evoked by Acinetobacter bacteria. Autoimmunity: concepts and diagnosis at the cutting edge. *Annals of the New York Academy of Sciences, 1050*, 417–428.

Georgsson, G., Sigurdarson, S., & Brown, P. (2006). Infectious agent of sheep scrapie may persist in the environment for at least 16 years. *Journal of General Virology, 87*, 3737–3740.

Goovaerts, P. (1997). *Geostatistics for natural resources evaluation*. New York, Oxford: Oxford University Press.

Gordon, I., Abdulla, E. M., Campbell, I. C., & Whatley, S. A. (1998). Phosmet induces up-regulation of surface levels of the cellular prion protein. *Neuroreport, 9*, 1391–1395.

Haase, B., Doherr, M. G., Seuberlich, T., Drogemuller, C., Dolf, G., Nicken, P., et al. (2007). PRNP promoter polymorphisms are associated with BSE susceptibility in Swiss and German cattle. *BMC Genetics, 8*, 15.

Hesketh, S., Sassoon, J., Knight, R., Hopkins, J., & Brown, D. R. (2007). Elevated manganese levels in blood and central nervous system before onset of clinical signs in scrapie and bovine spongiform encephalopathy. *Journal of Animal Science, 85*, 1596–1609.

Imrie, C. E., Korre, A., Munoz-Melendez, G., Thornton, I. & Durucan, S. (2008). Application of factorial kriging analysis to the FOREGS European topsoil geochemistry database. *Science of the Total Environment*. doi:10.1016/j.scitotenv.2007.12.012.

Isaaks, E. H., & Srivastava, R. M. (1989). *Applied geostatistics*. New York: Oxford University Press.

Jóhannesson, T., Gudmundsdóttir, K. B., Eiríksson, T., Kristinsson, J., & Sigurdarson, S. (2004). Selenium and GPX activity in blood samples from pregnant and non-pregnant ewes and selenium in hay on scrapie-free, scrapie-prone and scrapie-afflicted farms in Iceland. *Icelandic Agricultural Sciences, 16–17*, 3–13.

Juling, K., Schwarzenbacher, H., Williams, J. L., & Fries, R. (2006). A major genetic component of BSE susceptibility. *BMC Biology, 4*, 33.

Kabata-Pendias, A. (2001). *Trace elements in soils and plants*, 3rd edn. Florida: CRC Press.

Manuelidis, L. (2007). A 25 nm virion is the likely cause of transmissible spongiform encephalopathies. *Journal of Cellular Biochemistry, 100*, 897–915.

McBride, M. B. (2007). Trace metals and sulfur in soils and forage of a chronic wasting disease locus. *Environmental Chemistry, 4*, 134–139.

National CJD Surveillance Unit (2005). *Creutzfeld-Jakob disease surveillance in the UK: thirteenth annual report 2004*. Edinburgh: National CJD Surveillance Unit.

National CJD Surveillance Unit (2007). *Creutzfeld-Jakob disease surveillance in the UK: fifteenth annual report 2006*. Edinburgh: National CJD Surveillance Unit.

Prusiner, S. B. (1982). Novel proteinaceous infectious particles cause scrapie. *Science, 216*, 136–144.

Purdey, M. (1998). High-dose exposure to systemic phosmet insecticide modifies the phosphatidylinositol anchor on the prion protein: the origins of new variant transmissible spongiform encephalopathies? *Medical Hypotheses, 50*, 91–111.

Purdey, M. (2000). Ecosystems supporting clusters of sporadic TSEs demonstrate excesses of the radical-generating divalent cation manganese and deficiencies of antioxidant co-factors Co, Se, Fe Zn-does a foreign cation substitution at prion protein's Cu domain initiate TSE? *Medical Hypotheses, 54*, 278–306.

Ragnarsdottir, K. V., & Hawkins, D. P. (2006). Bioavailable copper and manganese in soils from Iceland and their relationship with scrapie occurrence in sheep. *Journal of Geochemical Exploration, 88*, 228–234.

Salminen, R., Tarvainen, T., Demetriades, A., Duris, M., Fordyce, F. M., Gregorauskiene, V., et al. (1998). *FOREGS geochemical mapping field manual*. Espoo: Geological Survey of Finland.

Salminen, R. (Chief-editor), Batista, M. J., Bidovec, M., Demetriades, A., De Vivo, B., De Vos, W., Duris,

M., Gilucis, A., Gregorauskiene, V., Halamic, J., Heitzmann, P., Lima, A., Jordan, G., Klaver, G., Klein, P., Lis, J., Locutura, J., Marsina, K., Mazreku, A., O'Connor, P. J., Olsson, S. Å., Ottersen, R. T., Petersell, V., Plant, J. A., Reeder, S., Salpeteur, I., Sandström, H., Siewers, U., Steenfelt, A. & Tarvainen, T. (2005). *Geochemical atlas of Europe, Part 1: Background information, methodology and maps*. Espoo: Geological Survey of Finland.

Sigurdarson, S. (2000). Scrapie or "rida" in Iceland. *Norsk Veterinarisk Tidskift, 112*, 408–413 (in Norwegian).

Stevenson, M. A., Morris, R. S., Lawson, A. B., Wilesmith, J. W., Ryan, J. B. M., & Jackson, R. (2005). Area-level risks for BSE in British cattle before and after the July 1998 meat and bone meal feed ban. *Preventative Veterinary Medicine, 69*, 129–144.

Tongue, S. C., Pfeiffer, D. U., & Wilesmith, J. W. (2006). Descriptive spatial analysis of scrapie-affected flocks in Great Britain between January 1993 and December 2002. *Veterinary Record, 159*, 165–170.

Wackernagel, H. (1998). *Multivariate geostatistics*, 2nd edn. Berlin: Springer-Verlag.

Wong, B. S., Chen, S. G., Colucci, M., Xie, Z., Pan, T., Lui, T., et al. (2001). Aberrant metal binding by prion protein in human prion disease. *Journal of Neurochemistry, 78*, 1400–1408.

Zeng, F., Watt, N. T., Walmsley, A. R., & Hooper, N. M. (2003). Tethering the N-terminus of the prion protein compromises the cellular response to oxidative stress. *Journal of Neurochemistry, 84*, 480–490.