26. Combining environmental and medical data sets to explore potential associations between environmental factors and health: policy implications for human health risk assessments

Jennifer McKinley,[1] Ulrich Ofterdinger,[1] Sherry Palmer,[1] Chloe Jackson,[1] Damian Fogarty[2] and Anna Gavin[3]

How to cite this chapter:

McKinley, J.M., Ofterdinger, U., Palmer, S., Jackson, C., Fogarty, D.G. and Gavin, A.T., 2016 'Combining environmental and medical data sets to explore potential associations between environmental factors and health: policy implications for human health risk assessments' in M.E. Young (ed.), *Unearthed: impacts of the Tellus surveys of the north of Ireland*. Dublin. Royal Irish Academy.

DOI: https://doi.org/10.7486/DRI.9s16dn03n

This review paper discusses how Tellus and Tellus Border soil and stream geochemistry data have been used to investigate the spatial relationship between certain diseases and levels of potentially harmful elements (PHEs) in soils and water. The research hypothesis is that long-term low-level oral exposure to PHEs via soil and water may result in cumulative exposures that may act as risk factors for progressive diseases, including cancer and chronic kidney disease. Some public policy implications for regional human health risk assessments, public health policy and education are also explored, including better integration of multiple data sets to enhance ongoing medical and social research.

Rationale

Several factors combine to make Northern Ireland an important exemplar for the UK and Ireland for advancing collaborative approaches to investigating potential associations between environmental factors and health. These are: a varied geology representative of the geology of the UK and Ireland; the availability of comprehensive geological, environmental and medical data sets; and a pioneering collaborative research culture between health professionals, practitioners and geoscientists.

[1] Queen's University Belfast.
[2] Belfast Health Trust.
[3] Northern Ireland Cancer Registry, Belfast.

The acquisition, quality assurance and maintenance of data are crucial to ensure an ethical and robust collaborative partnership. Key data sets in the Northern Irish context include detailed information on bedrock and superficial geology and the comprehensive Tellus soil and stream geochemistry data. In addition, the Northern Ireland Cancer Registry (NICR), the UK Renal Registry (UKRR) and the Northern Ireland Neighbourhood Information Service (NINIS) provide additional sources of anonymised population data. Combining geological, soil and medical data with demographic information offers the potential to examine the impact of specific environment factors on public health. The accuracy and potential complementarity of the data sets creates the opportunity for a multidimensional spatial investigation of a range of variables. These include: the influence of geographical distribution; total concentration of certain PHEs and their level of oral bioaccessibility; and the extent and variation of different diseases.

It is important also to consider the process by which the population may become exposed to environmental PHEs through the source–pathway–receptor linkage. Routes of exposure (inhalation, dermal contact and ingestion) and rates of uptake via such routes are spatially variable and have a strong influence on both the internalised and externalised uptake of PHEs. Rural Northern Ireland, as an area with traditionally little movement of people, provides greater potential for prolonged exposure at one location or area.

Any indications of spatial correlations and inferred links between harmful elements in soil and health have implications beyond Northern Ireland for human health risk assessments, and for strategic health and education policy development related to such potential risks. The data from the Tellus Border Survey and anonymised, population-based medical data (cancer and renal) offer the possibility for similar studies in the Republic of Ireland.

We describe the results of comparing the extent of spatial correlation between certain PHEs in soil and the occurrence of selected diseases in Northern Ireland. This work presents a partnership between the Belfast Health Trust nephrology (kidney medicine) research group, NICR and the School of Geography, Archaeology and Palaeoecology, Queen's University Belfast.

Expanding the narrative: combining multiple data sets

Cancer

Cancer remains the second most common cause of death in Ireland, after diseases of the circulatory system (CSO, 2014), with overall cancer incidence expected to increase by 20% between 2010 and 2020 and by 30% between 2010 and 2030 (NCR, 2011) mainly due to population ageing. On average, 11,861 cases of cancer (including non-melanoma (NM) skin cancer) were diagnosed in Northern Ireland each year between 2008 and 2012 (NCR, 2014). Correlating the spatial distributions of cancer and terrestrial PTEs was thought to be a potentially useful line of research.

Complete and reliable incidence information provided by cancer registries is key for use in comparability and disease-clustering studies. In Northern Ireland, between 1993 and 2003, 8640 new cancer cases were diagnosed every year. The Registry receives and collects electronic data information on all neoplasms diagnosed in Northern Ireland including NM skin cancers, and has an extensive programme of quality assurance of the registry data (see http://www.qub.ac.uk/research-centres/nicr/AboutUs/). It allocates area of residence based on postcode. All information is held securely and anonymised before analysis, The information contained in the NICR database is a uniquely valuable resource and makes Northern Ireland a prime location in which to undertake a multilayered spatial analysis of the potential links between potentially toxic elements in soil and different cancer diseases.

Renal disease

Increasing prevalence of chronic kidney disease (CKD) is resulting in a significant strain on health budgets. In the UK, the prevalence of CKD stages 3 to 5 is 8.5%, with an average cost of renal replacement therapy, including dialysis and kidney transplants, of £20,000 to £30,000 per patient per year, comprising 1–2% of the annual health budget (NHS, 2010; Lewis, 2012). Detection of CKD is increasing annually in Northern Ireland, from a raw prevalence of 30 per 1000 adults (2007) to 48 per 1000 adults (2014) (NINIS 2016). A similar level of prevalence in some counties of the Border Region of the Republic is reported by Stack *et al.* (2014)..

The UKRR reports on all patients starting on renal/kidney replacement therapy (RRT, needing dialysis or a kidney transplant), using a detailed data set electronically downloaded every quarter from all 72 UK renal units (Gilg *et al.*, 2012). Regional and centre-level variations in prevalence rates are observed in the UKRR that are not explained by the known factors of age, ethnicity or preceding diseases such as hypertension and diabetes. The unknown aetiology (or causes) in renal disease patients in Northern Ireland was 17.2% of the total incidence rates in 2010 (Gilg *et al.*, 2012). Arsenic, cadmium, lead and mercury have been identified as nephrotoxins on the human body through the dermal, inhalation and oral ingestion exposure pathways (Ekong *et al.*, 2006; Brewster, 2007; Soderland *et al.*, 2010). Exploring the spatial distribution of naturally occurring nephrotoxins in soils offers the potential for fresh insights into the observed geographic variation of long-term chronic disease.

Potentially harmful elements in soils

A key focus of research since the completion of the Tellus Survey has been to investigate the potential risk to human health associated with elevated total concentrations of heavy metals and metalloids in soils. Using shallow soil geochemical data in conjunction with supplementary bioaccessibilty testing of selected soil samples following the Unified BARGE Method (UBM), Barsby *et al.* (2012) found that for some soils hosting elevated

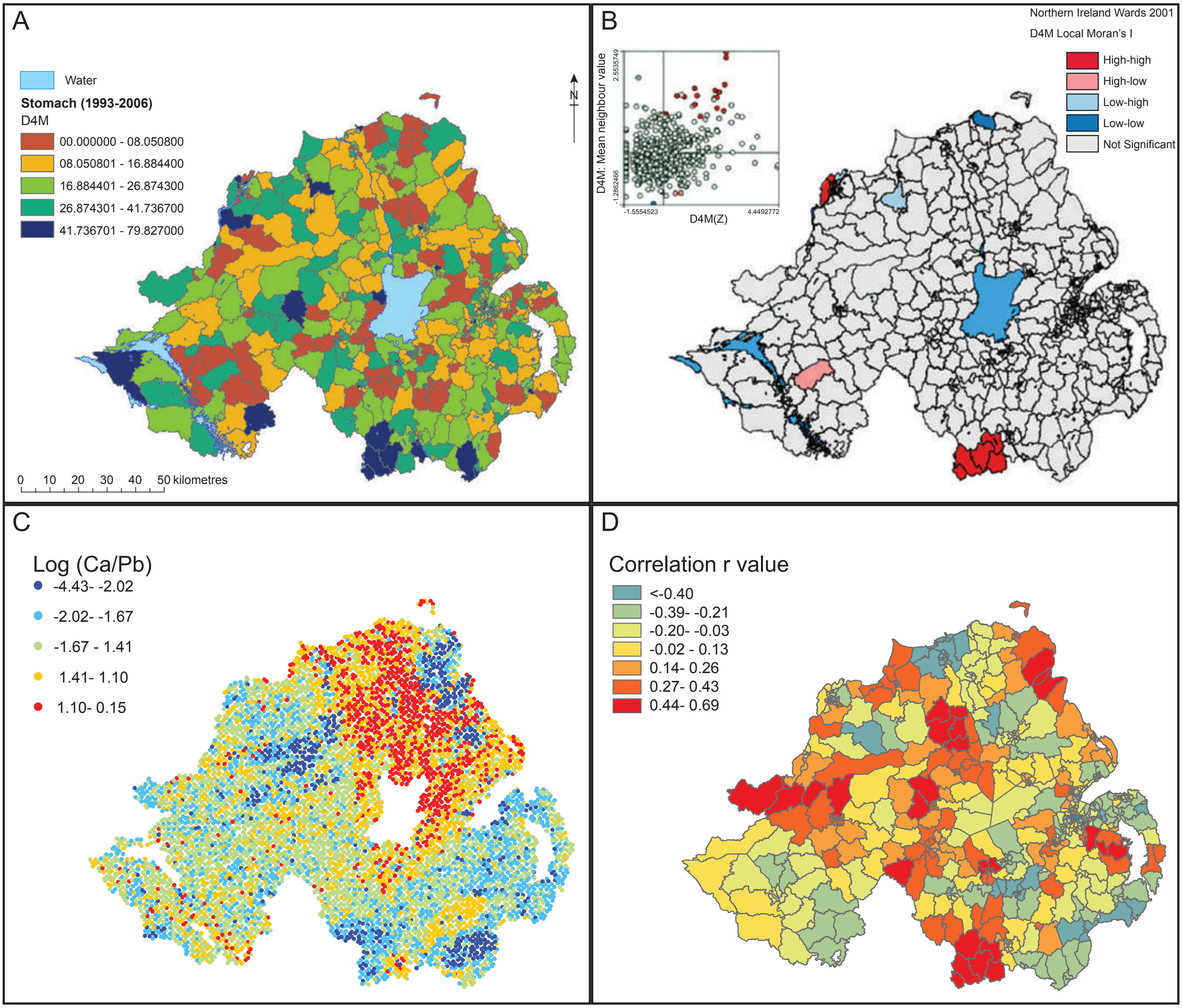

Figure 26.1. (A) ASIR mapping for male incidences of stomach cancer data (1993–2006); (B) Moran's I mapping for stomach cancer data (1993–2006); (C) log-ratio Ca/Pb map of total concentrations measured by XRF in soils; (D) GWR (correlation value) results for stomach cancer data (1993–2006) and arsenic (mg kg^{-1}) (A, B and D are adapted from McKinley *et al.*, 2013).

PHE concentrations, the measured oral bioaccessible fraction was relatively low. For other soil parent materials with relatively moderate total PHE concentrations, the measured oral bioaccessible fraction was relatively high. The discrepancy between high recorded total concentrations with lower oral bioaccessibility for certain areas has been related to elemental speciation and solubility both in the soil environment and during UBM extraction (Palmer *et al.*, 2014; Cox *et al.*, 2013). Palmer *et al.* (2015) found similarities in total and partial lead concentrations, suggesting that a large proportion of lead in Northern Ireland may be highly soluble and therefore potentially bioaccessible.

An integrated spatial methodology

Medical data

An interdisciplinary study by McKinley *et al.* (2013) investigated the incidence of 12 different cancer types (lung, stomach, leukaemia, oesophagus, colorectal, bladder, kidney, breast, mesothelioma, melanoma and NM basal and squamous skin cancers) over a 12 year period (1993 to 2006), with ethics approval from the NHS National Research Ethics Committee. The NICR provided anonymised cancer data sets in the form of postcode, gender, age group and disease code. In any study of medical data, population density and age distribution variations across the area of investigation need to be taken into account. Age standardisation takes the frequency of disease by age (and gender) group in each administrative unit and the standard population (Bland, 1996; Donnelly and Gavin, 2007) to provide the age standardised incidence rate (ASIR).

Mapped ASIRs were highest for NM squamous skin cancer for both sexes located in the south-east of Northern Ireland. Secondly, several wards in County Armagh showed high ASIRs for stomach cancer for the period 1993–2006 (Fig. 26.1a). High measured oral bioaccessibility shown for arsenic and lead align with the south Armagh region (Palmer *et al.*, Chapter 25, this volume). A local cluster technique known as Moran's I (McKinley *et al.*, 2013) identified statistically significant spatial outliers in the cancer disease data sets (Fig. 26.1b).

Geochemical data

The Tellus geochemical data sets are described by Young and Donald (2013) and by Young *et al.* (Chapter 3, this volume). The data are used to produce elemental concentration maps and to explore associations between elements. Fundamental in the use of geochemical data within a multidimensional spatial approach is an awareness that geochemical maps do not actually represent absolute abundances. At every sampled or interpolated point on an elemental map, a concentration value provides information only on the relative weight of one particular element to the total (Tolosana-Delgado and van den Boogaart, 2013). The challenge is how to respect this relative nature of geochemical data in the integration of multiple data sets. A log-ratio approach (Aitchison, 1986; Pawlowsky-Glahn and Buccianti, 2011) can be used to respect the constraints of the compositional geochemical data. This approach has the potential to shed light on the interactions between elements: the role of essential elements as protecting mechanisms against toxic elements, with others increasing the uptake of toxic elements as a result of similar absorption mechanisms. For example, a deficiency of calcium can enhance the uptake of Pb (Loghman-Adham, 1997). Several areas (Fig. 26.1c; for example the Mourne mountain areas) show a deficiency of Ca relative to Pb and therefore may indicate a higher uptake of Pb.

Exploring spatial relationships between disease data and explanatory factors

Care needs to be taken in comparing medical data aggregated to a certain spatial level to explanatory factors measured at regional or other levels of aggregation. This can lead to misconceptions of relations that are more related to the change of aggregation level than to the data. McKinley *et al.* (2013) used geographically weighted regression (GWR) to provide a local model of the relation between disease data and PHEs. Although the strongest relationship observed in this study was between stomach cancer and arsenic for wards in South Armagh, the results are complex (Fig. 26.1d). There may be a link with high levels of bioaccessible arsenic associated with the metasediments of the Gala Group of the Southern Uplands – Down–Longford Terrain (Barsby *et al.*, 2012). However, it is important to note that GWR does not establish a causal relationship and further multivariate data analysis is required within a multidimensional spatial framework.

Implications for policy: the need for greater dialogue between scientists and policy makers

The aim of interdisciplinary collaboration is to ensure that research findings are made available and any implications discussed between health professionals, practitioners and geoscientists. This ongoing partnership between geoscientists, NICR and the nephrology (kidney medicine) research group has raised a number of public policy implications for regional human health risk assessments and education. In addition this innovative interdisciplinary research strengthens the argument for better integration of multiple data sets to enhance ongoing medical and social research. In turn this underlines the need for greater dialogue between scientists and politicians.

Impact on the strategic public health agenda

This research shows that the complex geological history of Northern Ireland has resulted in geographic variation in elevated total concentrations and levels of bioaccessibility of some environmental toxins. Combined with observed and potentially linked geographic variation of long-term CKD and cancer, the research advocates the need for a more nuanced health policy that reflects this degree of variation across Northern Ireland. One implication from the findings is that screening priorities may need to change to adopt a targeted approach by health care professionals. For example, focused screening for certain types of cancer such as stomach cancer specifically in areas with elevated bioaccessible levels of arsenic could be considered. This has consequences for resource allocation to support early screening and detection, and poses the question of whether the current system has sufficient capacity to respond to this.

Avoiding a 'silo' approach to policy development

Information to increase awareness for health care professionals is key if the implications are to be understood and implemented. The scientific message that cumulative environmental exposures pose risk factors for progressive diseases including cancer and CKD needs to be communicated in an understandable way. This raises the issue of how to educate the next generation about the potential risks of PHEs and certain types of cancer or renal disease.

Enhancing opportunities for data linkage

This research has demonstrated the benefits of an integrated multilayered approach to multivariate data analysis facilitated by a cooperative partnership. However, it is clear that more work needs to be done in investigating environmental and health variables within a multidimensional spatial investigation. Opportunities include co-harvesting with other cross-discipline data sets, supporting integrative future planning of large spatial environmental surveys and longitudinal health studies. The main obstacles to integrating multiple data sources as part of mainstream research (such as cost and disparate methods of data capture) need to be identified. Enhanced opportunities for data linkage need to be encouraged and expedited to allow a multicomponent data analysis approach that will help generate even better insights to elucidate the relationship between our environment and health.

Acknowledgements

The NICR and GSNI are gratefully thanked for the use of the data sets. The NICR is funded by the Public Health Agency for Northern Ireland. The RIA Politics Plus programme is acknowledged for contributing to the rationale for this chapter. We are grateful to our reviewers, Mark Cave and Barry Smith.

References

Aitchison, J., 1986 *The Statistical Analysis of Compositional Data*. Monographs on Statistics and Applied Probability. London. Chapman and Hall. (Reprinted in 2003 with additional material by The Blackburn Press.)

Barsby, A., McKinley, J.M., Ofterdinger, U., Young, M., Cave, M.R. and Wragg, J., 2012 'Bioaccessibility of trace elements in soils in Northern Ireland', *Science of the Total Environment*, 433, 398–417. Available at http://nora.nerc.ac.uk/18896/. http://dx.doi.org/10.1016/j.scitotenv.2012.05.099.

Bland, M., 1996 *An Introduction to Medical Statistics*. Oxford. Oxford Medical Publications.

Brewster, U.C., 2007 'Chronic kidney disease from environmental and occupational toxins', *Connecticut State Medical Society*, 70, 4, 229–38.

Central Statistics Office (CSO), 2014 *The Statistical Yearbooks of Ireland*. Cork. CSO.

Cox, S.F., Chelliah, M., McKinley, J.M., Palmer, S., Ofterdinger, U., Cave, M.R., Wragg, J. and Young, M.E., 2013 'The importance of solid-phase distribution on the oral bioaccessibility of Ni and Cr in soils overlying Palaeogene basalt lavas, Northern Ireland', *Environmental Geochemistry and Health*, 35, 5, 553–67. Available at http://nora.nerc.ac.uk/502771/. http://dx.doi.org/10.1007%2Fs10653-013-9539-6.

Donnelly, D.W. and Gavin, A.T., 2007 'Survival of cancer patients in Northern Ireland: 1993–2004', in *Northern Ireland Cancer Registry*. Queen's University Belfast.

Ekong, E.B., Jaar, B.G. and Weaver, V.M., 2006 'Lead-related nephrotoxicity: a review of the epidemiologic evidence', *Kidney International,* 70, 12, 2074–84.

Gilg, J., Castledine, C. and Fogarty, D., 2012 'UK RRT incidence in 2010: national and centre-specific analyses', *Nephron Clinical Practice*, 120, Supplement 1, c1–27.

Lewis, R., 2012 *Understanding Chronic Kidney Disease: A Guide for the Non-specialist*. Keswick, UK. M&K Update.

Loghman-Adham, M., 1997 'Renal effects of environmental and occupational lead exposure', *Environmental Health Perspectives*, 105, 9, 928–38.

McKinley, J., Ofterdinger, U., Young, M.E., Gavin, A. and Barsby, A., 2013 'Investigating local relationships between trace elements in soils and cancer data', *Spatial Statistics*, 5, 25–41. Available at http://nora.nerc.ac.uk/502768/. http://dx.doi.org/10.1016/j.spasta.2013.05.001.

National Cancer Registry (NCR), 2011 *Annual Statistical Report.* NCR. Cork, Ireland. Available at http://www.ncri.ie/pubs/pubs.shtml.

National Cancer Registry (NCR), 2014 *Cancer in Ireland 1994–2012: Annual Report of the National Cancer Registry.* Cork. NCR.

National Health Service (NHS), 2010. *Kidney Disease Facts and Figures.* London. NHS.

Northern Ireland Cancer Registry (NICR), 2014 *Cancer Statistics by Local Government District: Patients Diagnosed in Northern Ireland during 2008–2012.* Belfast. NICR.

Northern Ireland Neighbourhood Information Service (NINIS), *2016 Disease Prevalence (Quality Outcomes Framework).* Available at http://www.ninis2.nisra.gov.uk/public/PivotGrid.aspx?ds=6944&lh=73&yn=2007-2015.

Palmer, S., Cox, S.F., McKinley, J.M. and Ofterdinger, U., 2014, 'Soil geochemical factors controlling the distribution and oral bioaccessibility of nickel, vanadium and chromium in soil', *Applied Geochemistry*, 51, 255–67.

Palmer, S., McIlwaine, R., Ofterdinger, U., Cox, S.F., McKinley, J.M., Doherty, R., Wragg, J. and Cave, M., 2015 'The effects of lead sources on oral bioaccessibility in soil and implications for contaminated land risk management', *Environmental Pollution*, 198, 3, 161–71. Available at http://nora.nerc.ac.uk/510763/. http://dx.doi.org/10.1016/j.envpol.2015.01.004.

Pawlowsky-Glahn, V. and Buccianti, A., 2011 *Compositional Data Analysis: Theory and Applications,* 12–28. Chichester, UK. Wiley.

Soderland, P., Lovekar, S. and Weiner, D.E., 2010 'Chronic kidney disease associated with environmental toxins and exposures'. *Advanced Chronic Kidney Disease*, 17, 3, 254–64.

Stack, A.G., Casserly, L.F., Cronin, C.J., Chernenko, T., Cullen, W., Hannigan, A. *et al.*, 2014 'Prevalence and variation of chronic kidney disease in the Irish health system: initial findings from the National Kidney Disease Surveillance Programme', *BMC Nephrology*, 2014, 15, 185. Available at http://bmcnephrol.biomedcentral.com/articles. http://dx.doi.org/10.1186/1471-2369-15-185.

Tolosana-Delgado, R. and van den Boogaart, K.G., 2013 'Joint consistent mapping of high dimensional geochemical surveys', *Mathematical Geosciences*, 45, 983–1004.

Young, M.E. and Donald, A., 2013 *A Guide to the Tellus Data.* Belfast. Geological Survey of Northern Ireland. Available at http://nora.nerc.ac.uk/509171/.

Table of Contents:

Unearthed: impacts of the Tellus surveys of the north of Ireland
First published in 2016 by the
Royal Irish Academy
19 Dawson Street
Dublin 2
www.ria.ie

ISBN: 978-1-908996-88-6

British Library Cataloguing-in-Publication Data. A catalogue record is available from the British Library.

Design: Alex Donald, Geological Survey of Northern Ireland.
Index: Brendan O’Brien.
Printed in Poland by L&C Printing Group.