Environment International 84 (2015) 90–93

Contents lists available at ScienceDirect

Environment International

journal homepage: www.elsevier.com/locate/envint

Assessing correlations between geological hazards and health outcomes: Addressing complexity in medical geology

Nicola Ann Wardrop [a,1], Jennifer Susan Le Blond [b]

[a] Geography and Environment, University of Southampton, Southampton, UK
[b] Department of Earth Science and Engineering, Imperial College London, London, UK

ARTICLE INFO

Article history:
Received 26 May 2015
Received in revised form 9 July 2015
Accepted 20 July 2015
Available online xxxx

Keywords:
Medical geology
Environmental epidemiology
Exposure assessment

ABSTRACT

Background: The field of medical geology addresses the relationships between exposure to specific geological characteristics and the development of a range of health problems: for example, long-term exposure to arsenic in drinking water can result in the development of skin conditions and cancers. While these relationships are well characterised for some examples, in others there is a lack of understanding of the specific geological component(s) triggering disease onset, necessitating further research.
Objectives: This paper aims to highlight several important complexities in geological exposures and the development of related diseases that can create difficulties in the linkage of exposure and health outcome data. Several suggested approaches to deal with these complexities are also suggested.
Discussion: Long-term exposure and lengthy latent periods are common characteristics of many diseases related to geological hazards. In combination with long- or short-distance migrations over an individual's life, daily or weekly movement patterns and small-scale spatial heterogeneity in geological characteristics, it becomes problematic to appropriately assign exposure measurements to individuals. The inclusion of supplementary methods, such as questionnaires, movement diaries or Global Positioning System (GPS) trackers can support medical geology studies by providing evidence for the most appropriate exposure measurement locations.
Conclusions: The complex and lengthy exposure–response pathways involved, small-distance spatial heterogeneity in environmental components and a range of other issues mean that interdisciplinary approaches to medical geology studies are necessary to provide robust evidence.

1. Introduction

The geological characteristics of the earth's surface can directly influence human health via the ingestion, inhalation or absorption of specific elements or compounds derived from naturally occurring materials (e.g. Davies et al., 2004; Skinner, 2007). The degree to which we understand the relationship between exposure and health outcomes, however, varies significantly between different geological hazards within the environment. For example, the relationship between exposure to water and food supplies contaminated with arsenic, and the development of skin conditions and a variety of cancers is well known (Bhattacharya et al., 2012; Naujokas et al., 2013). However, while the association between specific soil types and the development of podoconiosis (non-infectious elephantiasis) has been established, the specific components within the soil that may trigger the onset of podoconiosis have not yet been identified (Molla et al., 2014). When considering the discrepancies in our understanding of geological hazards, there are a number of important issues that must be addressed to enable us to explore the relationship between the environment and human health, most notably the compatibility between data collected to determine the potential hazard within the environment and that gathered to estimate disease occurrence.

Using statistical methods to link epidemiological data with geological characterisations can provide improved understanding of the etiologies of environmental diseases, but this linkage is not a straightforward one. Using a range of examples from medical geology, this paper aims to highlight several important complexities that need to be taken into account in research examining the relationships between geological hazards and health outcomes. A range of methodological approaches are discussed and evaluated which may allow these complexities to be addressed in future research.

2. Discussion

2.1. Characterising heterogeneity of geological variables

The aim of a geological survey is to map variability across a certain domain (sample area), providing a distribution of a variable or variables

---

E-mail address: Nicola.Wardrop@soton.ac.uk (N.A. Wardrop).
[1] Building 44, Highfield Campus, University of Southampton, University Road, Southampton, Hampshire SO17, 1BJ, UK.

http://dx.doi.org/10.1016/j.envint.2015.07.016

(i.e., concentration of metals in soil) in space and time. Essentially, a robust sample plan for the survey will reflect the purpose of the investigation, for example whether the map is to make local predictions across the domain, detect the presence/absence of certain components within the domain, or monitor whether the situation has changed over time (and space). When considering the vast number of exposure scenarios possible in the environment, within different environmental domains (e.g., air, soil/food and water) and via assorted routes (ingestion, absorption and inhalation), a broader perspective may need to be employed to identify the characteristics of the study area.

The traditional approach to map soil within a domain is to conduct a survey and collect soil samples for analysis, either in the field or in the laboratory, but sampling strategies are often defined by practical limitations such as funding constraints or logistical impracticalities. Geostatistical modelling methods (with or without the use of covariates), such as Kriging (a method for spatial interpolation), can be applied to investigate spatial variation in observations across the domain of interest, and importantly, to make use of this variation (spatial autocorrelation) to provide accurate spatial predictions at un-sampled locations. The distribution of soils will be determined by various environmental (e.g., parent rock type, climate, hydrology etc.) and anthropogenic (e.g., farming activities, pollution sources etc.) factors occurring at different spatial and temporal scales. In terms of spatial variation, targeted sampling is often compulsory due to the high cost of sample collection and analysis. If soil in the sampling area is highly variable (heterogeneous), the time needed to sample and costs of analyses will be high in order to obtain a sufficient spatial resolution to capture the variability (Vitharana et al., 2005).

When considering the contribution of certain environmental components within a health-related investigation, it is also crucial to incorporate temporal variation within the domain in order to more accurately estimate the exposure. In studies monitoring air pollution, for example particulate matter within a certain size range (e.g., $PM_{2.5}$ or PM_{10}), data for the particulate burden may be collected at point sources in the study area. This data can then be interpolated using other acquired variables (meteorological conditions, urban architecture, and information on the sources of particulate, for example motor vehicle movement) that will impact the density and distribution of the particulate matter over time and space. This information can be used to create maps, defining the variability of the hazard over the sample area. When used in conjunction with public health policy and exposure limits these outputs can be effective in identifying 'at risk' areas where the hazard is greatest.

2.2. Characterising heterogeneity of health outcomes

Epidemiological data can be either primary data (generated for the specific research purpose for which they are being used) or secondary data (generated for a purpose different from that for which they are being used, e.g. routine surveillance systems, or previous epidemiological studies) (Olsen, 2008; Woodward, 2013). The underlying population distribution and, therefore, the distribution of health outcomes are both inherently spatially heterogeneous, as are potential geological hazards. When considering the health impacts of geological exposures, it is clearly important to consider this spatial heterogeneity; thus, epidemiological data should have spatial attributes (Pfeiffer et al., 2008; Rothman et al., 2008).

Routine surveillance data will often include information on the administrative area in which individuals reside, allowing the aggregation of cases to specific administrative areas and the presentation of maps of case counts, or in combination with population data (e.g. from census data), prevalence or incidence (Beale et al., 2008; Lawson, 2006). The use of cross-sectional or cohort studies, in which health outcomes are assessed in individuals (rather than aggregates of individuals), gives greater opportunity to attach precise geographical locations, as geographical coordinates can be recorded for individuals' homes or alternative locations (Pfeiffer et al., 2008) and hence constrain exposure over time.

2.3. Linking geological hazards and health outcomes

The detection of unexpected health outcomes (often signified by unusually high incidence) in a population, suspected to be caused by exposure to a naturally occurring hazard, may instigate a geo-epidemiological study. Thus, acquisition of epidemiological data will typically be the initial response, followed by the collection of geological information to complement this dataset. The domain of interest needs to be considered from the outset as there is little point in assessing health outcomes in an area where the putative geological character does not vary. Thus, the study area should aim to encompass a range of values for the variables that can be measured to determine the hypothetical hazard. In addition, fundamental issues to consider are the potential mechanism of exposure (e.g. the environmental media in which the hazard exists and the route of exposure) and how the individual's exposure may vary within the population (e.g. genetic propensity, age, behaviour), both of which can be used to develop a dose–response relationship for the hazard.

To establish correlative relationships between the potential geological hazard and health outcomes, the two data sources (the epidemiological and geological surveys) need to be linked to allow statistical analysis. There are different ways of doing this. Where aggregated health outcome data are available within administrative units, data will be linked at the population level as in an ecological study (Woodward, 2013). This approach requires the environmental component(s) thought to be contributing to the disease to be collectively characterised within administrative areas, for example by calculating mean values for each area. Examining correlations in this way can be less demanding than for individual level studies (Nielsen and Jensen, 2005). However, within administrative units (often defined by political boundaries) the components within the environment contributing to the disease are likely to be highly variable and correlations detected at population level may not exist at individual level. Thus, these studies are useful for hypothesis generation for further study and can provide a useful means for the initial assessment of potential causative agents, but are prone to bias and the "ecological fallacy" (Morgenstern, 1982).

Epidemiological investigations at the individual level provide more detailed evidence of the correlations between environmental exposure and health outcomes, although the acquisition of suitable data is typically more time consuming and costly. Survey methods can be used to collect epidemiological data on health outcomes and exposures in individuals (e.g. case–control, cohort or similar study), but assigning quantitative measures of exposure to the environmental component to individuals is difficult. Ecological exposure data (e.g. mean values within an individual's area of residence) can be linked to individual level health outcome data, although this may not adequately capture heterogeneity in the environmental component, or individual level exposures (Hatch and Thomas, 1993; Nielsen and Jensen, 2005). Estimating exposure to the environmental component for each individual (e.g. at their home) allows us to directly link exposure and outcome information at an individual level, but is more challenging logistically and incurs greater financial costs (Hatch and Thomas, 1993). In addition, individual exposure estimates may be based on subjective information (e.g. questionnaire responses), with the potential to introduce measurement bias. Where it is not possible to take a physical measurement of hazard exposure for each individual included in the study, geostatistical methods may be beneficial. Geostatistical model-based predictions, such as Kriging, can be used to produce spatially continuous estimates of a value of interest (e.g. concentrations of the environmental component associated with the disease) based on an even coverage of data from the sample area: the spatially continuous estimates can then be used to provide exposure estimates for individuals based on their spatial locations (Goovaerts, 2014).