Soil and human health: an epidemiological review

R. L. Hough

The Macaulay Institute, Craigiebuckler, Aberdeen AB15 8QH, UK

Summary

Two different approaches have been used to study relationships between soil and human health: aggregate- and individual-level. Aggregate-level is the primary approach used in the geosciences and broadly relates spatial soil characteristics to geographic incidence of disease. However, this may not be appropriate for a wide range of exposures and disease outcomes. For example, many diseases with long latency periods are associated with early life or cumulative exposure to the causative agent, rather than an individual’s current geographical location. Public health scientists and epidemiologists often refer to aggregate-level studies as being ‘hypothesis forming’, and consider individual-level approaches as the ‘gold standard’ for investigating causes of human health outcomes. This paper investigates the appropriateness of individual-level study for disease outcomes associated with soil by reviewing the weight of evidence from individual-level studies that have included exposure to soil as a risk-factor for disease. The majority of these studies are very specific, inspired by explicit case reports and medical records. The review showed that exposure to soil was implicated mainly in the spread of enteric parasites, but also in the incidence of certain cancers, bacterial infections, and mycoses. There was little evidence for specific soil contaminants as causative agents of disease at an individual level but there was also little evidence contrary to this. Further individual-level studies into soil:human health scenarios are required in order to derive more appropriate dose-effect relationships for regulatory science and risk assessment.

Introduction

Worldwide, environmental and occupational policies increasingly are based on Quantitative Risk Assessment (QRA), a technique designed to forecast risks that cannot be measured directly (Wartenberg & Chess, 1997). The EU and USA are leading proponents of the use of QRA for the assessment of risks posed by contaminated soils to human health (DETR 2000; Holgate 2000). It is generally recognized that human, epidemiological data, if available are preferred as the starting point for QRA above the use of data from animal studies (Goldbohm et al., 2006). Important risk assessment models, such as the CLEA (Contaminated Land Exposure Assessment; DEFRA and Environment Agency 2002) model (implemented in the UK as a ‘supporting initiative’ to contaminated land redevelopment) have been developed almost exclusively on animal data (DEFRA & Environment Agency 2002). Human health outcomes associated with contaminated soil are considered chronic or subchronic and therefore may be difficult to study. In fact, there is a general consensus amongst regulatory scientists in the UK that it is virtually impossible to untangle causal relationships from a multitude of confounding and modifying factors associated with the contaminated soil environment (Bardos 2003). This opinion is very much structured by single contaminant (or causal factor)-multiple outcome relationships, specifically associated with the approach and contaminants covered by the CLEA model. Human epidemiology tends to look at health issues from the perspective of the outcome, i.e. single outcome-multiple causal factors. This approach may enable associations to be defined through weight of evidence.

Broadly speaking, two different approaches have been used to study the potential affects of soil on human health: aggregate- and individual-level. In aggregate-level approaches, the incidence of a disease outcome across a geographical area is related statistically to selected soil characteristics. For example, it may be found that school examination statistics for populations living in areas with elevated concentrations of lead in the soil, report a lower than average level of attainment.

The individual approach attempts to relate the outcome (poor examination results) to the exposure of interest (lead), at an individual level, i.e. a number of school children will be recruited who show either good or poor aptitude in recent examinations. For each child, some measurement of exposure to lead in the soil will be determined, and potential exposure to other sources of lead will also be determined; other factors that might cause this
particular child to perform poorly during examination will be documented and/or quantified (these factors will be 'controlled for' during statistical analysis). Finally, the probability of being exposed to environmental lead and underperformance during examination will be compared to the probability of not being exposed to lead and underperformance during examination, i.e. determine the effect of lead exposure on risk of underperformance.

This review aims to answer the question of whether there is any epidemiological evidence (including weight of evidence) to suggest that agents within soil have an adverse effect on human health. Therefore, this review does not include toxicological weight of evidence, or aggregate-level studies (termed 'ecological' in epidemiology nomenclature) that broadly relate spatial variation in soil parameters to incidence of disease, the latter having been covered comprehensively by Oliver (1997). Ecological-level studies are often described as being 'hypothesis-forming' among public health scientists. This is mainly because, at best, they can only provide a crude attempt to separate the exposure of interest from numerous other causal, confounding, and modifying variables. Typically, exposure to soil occurs in poorly controlled environments. This means it is often difficult to 'untangle' soil-related variables from other environmental and occupational exposures. Long latency periods of certain diseases also mean that current environmental data may not be relevant to the outcome of interest. Ecological studies have the advantage that they are relatively quick (and inexpensive) and are therefore attractive to policy-makers. For the epidemiologist they provide indication of where research efforts should be focussed.

Due to these reasons, the focus of this review is individual-level studies that have investigated exposure to soil and/or soil-related activities as risk factors for specific disease outcomes. Studies that have shown exposure to soil as a risk factor for specific biomarkers of disease, e.g. lead in blood, have been omitted as these do not point to specific burden of disease. As a result, this review is structured by disease outcome, rather than by soil characteristics. The objective was to determine which 'soil' exposure indices have been used in epidemiological studies, and identify areas (or health outcomes) where further epidemiological evidence is needed.

Methodology

Search strategy and study evaluation

This was a purposive review of peer-reviewed literature related to soil exposure as a risk factor for disease. In a few cases, this was supplemented with literature from web-accessible documents and other grey literature. Computerized database searching of a range of international databases was carried out: MEDLINE (MEDLINE Database, National Library of Medicine, Bethesda, MD, USA www.biomednet.com), BIDS (Joint Information Systems Committee, University of Bath, Bath, UK, www.bids.ac.uk), PubMed (National Library of Medicine, http://www.ncbi.nlm.nih.gov/PubMed/) and Web of Science.

A systematic, staged search strategy was employed using the following search terms: 'Epidemiology' OR 'risk factor' OR 'exposure' AND 'soil' OR 'geochemistry' OR 'environment'.

Over 2600 references were found. The search was further refined to 394 papers that investigated soil as a risk to health. Only papers written in English were included. Original full texts were obtained for all studies. Upon examination of the full texts, the following types of studies were excluded (except where a method was used that showed principles relevant to the context of this review): aggregate- or ecological-level designs, descriptive and surveillance studies, and other papers that only made inference to soil being a risk factor for disease. The remaining 32 studies that quantified the level of risk to individuals from soil or soil-related activities (e.g. gardening) formed the basis of this review, although other papers within the database were used as background material. Those studies that provided the greatest weight of evidence according to the methodology described by Goldbohm et al. (2006) were considered in most detail to illustrate epidemiologic approaches.

Individual-level approach

An overview of epidemiologic approaches can be found in Gerstman (2003). Individual-level studies can be either prospective or retrospective. In a prospective study, the health status of individuals in a cohort(s) is monitored over time. Prospective study is very expensive and time consuming, with studies that look at rare outcomes lasting for decades. Instead, many studies opt for a retrospective approach, with the case-control design considered the 'gold standard'. Under this study design, new cases (individuals with the outcome of interest) are selected and matched to controls. New cases are confirmed, usually with some form of biological test. The number of cases and controls required will depend on the expected incidence of the disease. In some studies, a number of different control groups are matched to the cases. Information on individuals' exposure (including specific risk factors), occupational, health and dietary histories are obtained, usually through interviews or questionnaire approaches. Data on other factors that might confound or modify the exposure-outcome relationship of interest are also collected. Finally, a comparison is made between the probability of being exposed and also being a case, and the probability of being exposed and being a control. This comparison is typically performed using a logistic regression model that may be univariate, or multivariate, the latter being used to account for (adjust for) the confounding and modifying variables identified.

Minimization of bias is imperative in epidemiologic study, as the method of selection of cases and controls can introduce bias. For example, where both cases and controls are drawn from hospital registers, there is the assumption that this sample truly reflects the population that they represent. Hospital-based
controls should have diseases that are not associated with the outcome of interest and selecting multiple control groups can act as insurance against this problem. Other sources of bias include recall bias by individuals. These issues are tackled using well-designed questionnaires with, where possible, previously validated questions.

Significance in epidemiologic study

Typically, the significance of any exposure–outcome association is reported with the ‘odds ratio’ (OR). This is usually the ratio of the ‘odds’ (probability) of being exposed to an agent and of developing the disease of interest, to the ‘odds’ of being non-exposed and developing the disease of interest. The OR is therefore a relative scale, with any value exceeding unity indicative of elevated risk. Confidence in the estimation of OR is usually expressed by 95% confidence intervals (CI). The wider the confidence intervals, the less the certainty that can be placed on the estimate of OR. Where the confidence intervals include one, the association is considered not statistically significant ($P \geq 0.05$).

Significance is also quantified by rate ratios (RR), which are the ratio of the incidence rate of disease in the exposed group to the incidence rate in the non-exposed group. Incidence rates are normalized by person years, so for rare outcomes (e.g. cancer) they can be considered equal to the OR. However, when a disease is not rare, RR tends to be greater than OR.

Cancer

Cancer is a very important disease outcome, not only in terms of the impact that it has on the people affected, but also on the economics of health care provision. Many cancers are difficult and expensive to treat and have therefore become the target of Governmental and public health campaigns (Boffetta, 2004). Like many rare diseases (i.e. low incidence rate), cancer presents a number of difficulties to those who want to study its causes. Being rare, researchers require large populations to attain the required level of statistical power to ‘see’ associations with environmental variables. Also, because cancer typically has a long latency period (30–40 years), retrospective information on patients’ exposure to risk factors throughout their lifetimes is needed.

Soil properties have been associated at an ecological level with a range of different cancers (Boffetta, 2004) including: (i) malignant pleural mesothelioma caused by fibrous zeolite (Sahin et al., 1993); (ii) lung cancer and other childhood cancers such as leukaemia caused by radon gas (Peake & Hess, 1987; Henshaw et al., 1990); (iii) gastrointestinal cancer associated with salinity of soil (Memik et al., 1996); and (iv) oesophageal cancer associated with low fertility soils (Laker et al., 1981). However, only a limited number of studies actually relate individual-level exposure to soil as a risk-factor for cancer.

Kaposi’s sarcoma

Ziegler (1993) hypothesized a relation between African (or endemic) Kaposi’s sarcoma (a cancer that develops in connective tissues such as cartilage, bone, fat, muscle, blood vessels, or fibrous tissues) and the presence of volcanic soil. In areas of high incidence, patients displayed immune suppression of the extremities (hands, feet). Zeigler put forward the theory that barefoot contact with the volcanic soil is a causative mechanism (Ziegler, 1993; Ziegler & Katongole-Mbidde, 1996; Ziegler et al., 2001). Kaposi’s sarcoma has also been described frequently in southern Italy with high incidence rates reported in the registry of Ragusa, Sicily (3.00 per 100 000 men and 0.50 per 100 000 women standardized to the Italian population) (Franceschi & Geddes, 1995). An ecological study of 139 cases of Kaposi’s sarcoma in Sardina (1987–93) in relation to seven different soil types did not show excess incidence in areas of siliceous volcanic soils (Montesu et al., 1995). However, the RRs described in this study suffered from very wide confidence intervals.

Montella et al. (1996) conducted a case-control study in the Campania region of Italy that included the densely populated flanks of Mount Vesuvius. The study consisted of 70 histologically confirmed cases and 168 controls. Cases were defined as being resident in Campania and had been referred to three major hospitals in Naples between 1980 and 1993. AIDS-related Kaposi’s sarcoma was ruled out by means of anonymous linkage to the mandatory Italian AIDS register. Controls were matched by sex, age (± 2 years), year of diagnosis, and hospital. Control diagnosis included benign or malignant tumours (chiefly epithelial cancers), skin diseases (eczema, psoriasis, urticaria), and various infections. Information on the individuals’ place of birth and residence was obtained from medical records. The presence of basaltic lava was determined from maps of the last eruption in 1944 (Cortini & Scandone, 1987) and was used as a proxy for exposure to volcanic soil. The data were analysed by an unconditional multiple logistic regression model adjusted for sex, age, and hospital. Montella et al. (1996) found individuals were more than twice as likely to develop Kaposi’s sarcoma if they were born in a volcanic area (OR = 2.12; 95% CI 1.04–4.31), which suggests that early life exposure to volcanic soils is an important risk factor. Current exposure did result in elevated risk (OR = 1.84); however, the confidence interval included 1 (95% CI 0.85–3.98), indicating that this was only a weakly significant association. These results suggest that indices of current exposure might not be relevant to the risk of developing diseases with a long latency period, such as cancer. However, it should be noted that the presence of lava is a very crude proxy of exposure and does not indicate contact time with bare skin (hypothesis put forward by Ziegler, 1993; Ziegler & Katongole-Mbidde, 1996; Ziegler et al., 2001). Also, the sample number in this study was very small when considering the expected incidence rate reported in Franceschi & Geddes (1995). The use of a
hospital-based methodology also has the potential to introduce selection bias (Fletcher et al., 2004a).

Lung cancer

Experimental work has shown that some crops (including rice, sugar cane, wheat and millet) take up silica in the form of soluble monosilic acid (H$_4$SiO$_4$) from the soil and deposit it in the epidermal cells of the leaves (Fox et al., 1969). The deposited biogenic silica is amorphous and can exist in the form of fibres similar in morphology to asbestos, and might therefore cause similar health effects (Newman, 1983; Rabovsky, 1995).

Amre et al. (1999) looked at risk factors for lung cancer among sugar cane farmers and cane processors in India (Maharashtra) – an area with considerable sugar production and a large agricultural population. The study looked at both farmers and sugar cane process workers. Amre et al. (1999) recruited newly diagnosed, histologically confirmed, cases (n = 128) of lung cancer between 1996 and 1998 and controls (n = 298) from six major cancer treatment centres within the province. Controls diagnosed with other cancers were matched to cases by age (± 10 years), sex, district of residence, and timing of diagnosis (± 2 months of that of the case). Use of conditional logistic regression accounted for the matching in the study design and revealed: an elevated yet insignificant risk associated with harvesting crops (OR 1.41; 95% CI 0.70–2.90); an increased risk for those who worked on cane farms (OR = 1.92; 95% CI 1.08–3.40); and an increased risk for those involved in crop preparation (OR = 1.81; 95% CI 0.99–3.27). The risk also increased (by an OR of 1.21, 95% CI 1.20–1.40) for every 10 years worked. Those involved in farm preparation for > 1160 days were over twice as likely to develop cancer than those who had worked < 1160 days (OR = 2.3; 95% CI 1.10–4.70).

The study by Amre et al. (1999) was the only one that looked at individual risks associated with soil related activities, but it was a hospital-based design and therefore susceptible to selection bias. Rothschild & Mulvey (1982) reported a more than twofold increase in risk of lung cancer (OR = 2.40; 95% CI 1.70–3.60) in Louisiana sugar cane farmers after (statistically) controlling for the effects of smoking and exposure to asbestos. However, risks within specific activities were not assessed. Brooks et al. (1992) found an increase in risk of lung cancer in sugar cane workers in Florida (OR = 1.80; 95% CI 0.50–7.50). However, the confidence intervals suggest that this association was not significant. A similar result was reported by Miller et al. (1993), but this study also did not look at specific soil-related activities.

Leukaemia

A descriptive incidence study (Kolb et al., 1993) showed rates of leukaemias to be elevated in the town of Stadtallendorf, Germany. During World War II, two armament factories were active in the town and uncontrolled dumping of armament wastes was commonplace leading to elevated concentrations of toluene in the soil (Kilian et al., 2001). A case-control study to investigate the risk factors associated with this cluster of disease was conducted (Kilian et al., 2001). Cases (n = 43) were identified within areas predominantly affected by the contamination. Cases were defined as individuals who had been diagnosed with either acute myelogenous leukaemia (AML) or chronic myelogenous leukaemia (CML) between the years 1979 and 1993. Controls (n = 123) were matched to cases by age and sex and were selected at random from the population register for 1994. Although a number of different soil-related exposure indices were investigated (ownership of a garden, gardening, kitchen-garden), only individuals who had been involved in growing their own food in their garden for greater than 10 years showed elevated risk (OR = 2.40; 95% CI 0.80–7.30). However, this association was only weakly significant as indicated by the confidence intervals that include 1.00. The study itself suffers from poor statistical power when considering the incidence rates reported in Kolb et al. (1993). Also, the selection of controls from only the contaminated areas might introduce bias into the analysis. Again, with a long latency disease, reconstructing exposure histories might have provided stronger risk factor–outcome associations (Fletcher et al. 2004b).

Deficiencies

For centuries, soil has been linked to a number of nutritional deficiencies. For example, Oliver (1997) reported that Avicenna (AD 980–1037) hypothesized that geophagy was a causal factor of anaemia. The link between geophagy and anaemia has been well-documented in developing countries (Geissler et al., 1998; Nchito et al., 2004) Soil-related deficiencies of trace elements such as selenium, copper, iron and zinc have also been implicated as causal factors for anaemia (Prasad et al., 1961; Okcuoğlu et al., 1966; Halsted, 1968; Minnich et al., 1968; Låg, 1987). Similarly, geophagy has been associated with hypokalemic myopathy (gastrointestinal wastage of potassium) (Mengel et al., 1962; Severance et al., 1988; Gonzalez et al., 1982; Chaushev et al., 2003).

Anaemia

Although there have been numerous studies that make the link between geophagy and anaemia, very few actually quantify risk at an individual-level. Antelman et al. (2000) investigated potential risk factors for anaemia among pregnant women with HIV in Tanzania. The reasons for such a specific case-group included: anaemia having been shown to be an important risk factor for poor pregnancy outcomes (Allen, 1997) and that infection with HIV during pregnancy has been associated with an increased risk of anaemia-related maternal death in developing countries (McDermott et al., 1996). The study by Antelman et al. (2000) was based on data from a large
(n = 1083), on-going randomized, controlled, double-blind clinical trial described by Fawzi et al. (1998). ‘Anaemia’ was defined as measured haemoglobin (Hb) < 110 g l⁻¹; with ‘severe anaemia’ defined as Hb < 85.0 g l⁻¹. The prevalence of reported geophagy during pregnancy was 28% (Antelman et al., 2000). Adjusted risk factor associations showed an increased risk of developing ‘anaemia’ associated with geophagy among HIV-infected pregnant women (OR = 1.67; 95% CI 1.02–2.74). The association with ‘severe anaemia’ was even stronger, suggesting that women who ingest soil are two and a half times more likely to develop ‘severe anaemia’ than those who did not partake in geophagia (OR = 2.47; 95% CI 1.66–3.69). These findings are supported by Desai et al. (2002), who recommended that geophagy should be considered as an additional indicator of severe anaemia.

Iodine deficiency disorders (IDD)

If the amount of utilizable iodine reaching the thyroid gland is inadequate, or if the thyroid gland is not functioning as it should, then hormone production is reduced, resulting in a group of conditions collectively referred to as iodine deficiency disorders (IDD) (Fernando et al., 1989). Iodine deficiency disorders have been described as ‘the world’s single most significant cause of preventable brain damage and mental retardation’ (El-Sayed et al., 1998). Goitre, the most common manifestation of iodine deficiency, is an enlargement of the thyroid gland in an attempt to make more efficient use of an inadequate iodine supply (Fordyce et al., 2000). A number of studies carried out in Sri Lanka since 1947 have suggested a link between goitre and soils with low concentrations of (bioavailable) iodine (Mahadeva et al., 1968; Deo & Subramanian, 1971; Fernando et al., 1989). However, the link to soil is only inferred in these studies, and not quantified at an individual level.

El-Sayed et al. (1998) carried out a cross-sectional study to investigate IDD among schoolchildren in Upper Egypt. Using a two-stage cluster-sampling technique, 6750 schoolchildren aged 8–10 years were screened for IDD through clinical examination of the thyroid gland and determination of urinary iodine. Iodine was also determined in Relevant soil and water samples were analysed for iodine. The prevalence rate of goitre was found to be 34.6%. Multiple logistic regression analysis showed that certain groups of schoolchildren were more likely to develop goitre. This included children living in places where the iodine content of the soil was less than 0.002 μg g⁻¹ (OR = 2.67; 95% CI 2.30–3.10).

Parasites

The literature on soil-transmitted parasites is vast, with a number of recent comprehensive reviews covering transmission, global public health impact, prevention and control (De Silva et al., 2003; Norhayati et al., 2003; Cappello, 2004; Savioli & Albionico, 2004; Savioli et al., 2004). However, there has been only a limited number of studies that quantify the association between exposure to soil and incidence of soil-transmitted parasites at an individual-level.

Cyclosporosis

Outbreaks of infection with the coccidian parasite Cyclospora cayetanensis in the United States of America during 1996 and 1997 led to a series of studies, both in the USA and in Guatemala. The outbreaks in the USA involved more than 1400 persons, and the source was thought to be raspberries imported from Guatemala (Herwaldt et al., 1997). Prior to these outbreaks, cyclosporosis was generally associated with children living in developing countries under poor sanitary conditions (Ortega et al., 1993; Hoge et al., 1995). Koumans et al. (1998) carried out a retrospective cohort study and a matched case-control study on the affected US population. The case-control study involved 24 laboratory-confirmed sporadic cases and 69 controls who were friends and/or acquaintances of the cases. The retrospective cohort study was conducted among two groups of persons who may have been exposed to Cyclospora during two separate social events where Guatemalan raspberries had been consumed.

Univariate logistic regression identified eating raspberries as the most significant risk factor in Cyclosporosis development for the case-control study, although the confidence intervals were very wide (OR = 6.00; 95% CI 1.10–31.7). Interestingly, other important risk factors included bare-skin contact with soil within 2 weeks before the onset of symptoms (OR = 5.40; 95% CI 1.40–20.7), and gardening within 2 weeks before the onset of symptoms (OR = 3.50; 95% CI 1.10–11.6). In a further multivariate analysis, bare skin contact with soil remained the only significant risk factor. Those who had soil contact within 2 weeks prior to the onset of symptoms were seven times more likely to develop Cyclosporosis than those who had not (OR = 7.20; 95% CI 1.50–33.8). Contact with soil was not investigated as a risk factor within the cohort study.

Risk factors for Cyclosporosis in Guatemalan raspberry-producing regions were also investigated (Bern et al., 1999). Here, a case-control study was conducted in 1997. Cases were recruited from health care facilities, with controls matched from individuals who had had a stool sample screened in a health-care facility during the period when cases were being recruited. Similarly to the study in the USA, contact with soil was shown to double the risk of developing Cyclosporosis (OR = 2.30; 95% CI 1.10–4.90). Overall, the studies suggest that contact with soil is the primary risk factor. However, eating produce that has come into contact with contaminated soil (e.g. raspberries) is also an important exposure pathway.

Toxoplasmosis

Toxoplasma, Listeria and Salmonella are the three most important pathogens carried by food in terms of illness and death in
the USA and Europe (Dubey, 2000). Toxoplasmosis is caused by infection with the protozoan parasite *Toxoplasma gondii*. Acute infections in pregnant women can be transmitted to the foetus and cause severe illnesses such as mental retardation, blindness, and epilepsy. An estimated 400–4000 cases of congenital toxoplasmosis occur each year in the USA (Hughes et al., 2000), and between 1 and 10 in 10 000 newborn babies across the EU (Gilbert, 1999). Cook et al. (2000) conducted a European multicentre (Naples, Lausanne, Copenhagen, Oslo, Brussels, and Milan) case-control study to identify sources of *Toxoplasma* infection in pregnant women. Cases were defined as pregnant women diagnosed with acute *Toxoplasma* infection between January 1994 and June 1995. The controls were the next four women who tested negative (in the same laboratory by the same screening methods as the cases) for the presence of antibodies specific for *Toxoplasma*. Only women from within the populations of the six countries were included in the study. Data were analysed with a multiplicative, unconditional logistic regression model that allowed for centre, maternal age, and interval between diagnosis and interview, to examine the risk associated with each exposure. A second multivariate, multiplicative, logistic regression model was performed with those variables with a *P*-value ≤ 0.25 in the first analysis. Although the eating of raw or undercooked meats was the strongest predictor, the univariate analysis highlighted contact with soil (OR = 1.90; 95% CI 1.40–2.80; *P* < 0.001) and living on a farm (OR = 1.60; 95% CI 0.90–2.70; *P* = 0.01) as being important risk factors. Contact with soil continued to display a strong association with risk of toxoplasmosis within the multivariate model (OR = 1.81; 95% CI 1.20–2.70; *P* = 0.005).

Jones et al. (2001) described a large cross-sectional study conducted in the USA. The Third National Health and Nutrition Examination Survey (NHANES III) was conducted between 1988 and 1994 by the National Centre for Health Statistics of the Centres for Disease Control and Prevention (NCHS, 1992, 1994). It was designed to obtain nationally representative statistics on a variety of health measures and conditions through household interviews, standardized physical examinations, and collection of blood samples in mobile examination centres (NCHS, 1992). The study included 27 145 individuals, of which 17 658 had been sera-tested for *T. gondii*. Analysis of these data indicated that employment in an occupation that involved exposure to soil increased the risk of developing toxoplasmosis (OR = 1.40; 95% CI 1.06–1.85). Bobić et al. (1998) reported a cross-sectional study that looked at risk factors for toxoplasmosis in reproductive age females (*n* = 1147) living in Belgrade. As found in the study of Cook et al. (2000), the most important risk factor was the consumption of undercooked meats. However, in women below age 20, exposure to soil through farming and/or gardening was significantly associated with infection (OR = 1.38; 95% CI 1.12–1.97; *P* = 0.037). Across all three studies, the increased risk associated with soil contact (1.38–1.60) was very consistent given the diverse populations.

**Nematode infections**

The prevalence of infection with helminth parasites is approximately two billion people worldwide (Crompton, 2001), which results in an estimated two million clinical episodes and 60 000 deaths annually (Walsh, 1990). Intestinal parasitism is common among children in developing countries (Holland et al., 1989; Gbakima et al., 1994; Akufongwe et al., 1995), and is the second leading cause of mortality in children (< 6 years of age) in rural Africa (Duconge, 1993; Gyorkos et al., 1996; both cited in Glickman et al., 1999). Little is known, however, about risk factors for infection (Glickman et al., 1999; Nishiura et al., 2002).

Toxocariasis is an infection predominantly caused by migration of the larvae of the roundworm *Toxocara canis* to organs and tissues. Humans can be infected either by ingestion of parasite eggs, or by contact with the teats of female dogs that have recently given birth. Childhood infections with *T. canis* have been associated with a number of serious health outcomes including endomyocarditis (De Cock et al., 1998), asthma (Chan et al., 2001), and epilepsy (Nicoletti et al., 2002). Iddawela et al. (2003) conducted a cross-sectional study to investigate risk factors for *T. canis* infection in 1020 children aged 1–12 years in Sri Lanka. Unconditional logistic regression showed 7- to 9-year-olds to be at greatest risk (OR = 3.08; 95% CI 1.95–4.87), with geophagia a very significant risk factor (OR = 6.37; 95% CI 3.87–10.5). Fan et al. (2004) conducted a cross-sectional study to determine risk factors for *T. canis* infection in Taiwan. The study focussed on the Bunun people, an aboriginal tribe living in mountainous (100 m to 600 m above sea level) regions in the east of the country. Twelve primary schools were identified in five different villages with a total of 618 children. Multivariate analysis showed that playing with soil was the primary risk factor and increased risk of *T. canis* infection by two and a half times (OR = 2.52; 95% CI 1.49–4.25; *P* < 0.001). Raising dogs was shown to be the second most important factor (OR = 1.83; 95% CI 1.04–3.19; *P* = 0.03). Children who habitually washed their hands almost halved their risk of *T. canis* infection (OR = 0.57; 95% CI 0.33–0.97; *P* = 0.04).

A study of the prevalence and intensity of *Ascaris lumbricoides* (usually asymptomatic, but can cause inflammation, fever and diarrhoea) in 492 children from five rural villages in the Northern Area (‘Baltistan District’) of Pakistan was conducted by Nishiura et al. (2002). Baltistan is a remote mountainous region with a stable population involved in subsistence-level farming. The populations (*n* = 300–600) of the five villages selected for the study were endemic for *A. lumbricoides*. Faecal samples were collected from participants. The study defined the presence of 0–4999 *A. lumbricoides* eggs per gram of faeces (EPG) as a ‘light
infection”; between 5000 and 50 000 EPG as a ‘moderate infection’; and > 50 000 EPG as a ‘heavy infection’. Univariate analysis showed that, along with location of house and hand washing after defecation, soil eating was a major risk factor in *A. lumbricoides* infection (*P* < 0.01). However, multivariate analysis did not confirm these findings, with soil eating the fourth most important risk factor (*P* = 0.28).

A very similar study was performed by Glickman *et al.* (1999) in rural Guinea, Africa. A stool examination was performed on 286 randomly selected children aged 1–18 years from three rural villages (Konindu, Dandakara and Dramela), population = 5360. Although this region was selected because it was the birthplace of one of the authors, it was also considered representative of this region of Guinea. Logistic regression was used to examine the relationship between risk factors and prevalence of *A. lumbricoides* infection and also hookworm (*Strongyloides stercoralis*). Results showed that gender (*P* = 0.01) and geophagia (*P* = 0.01) were the most important risk factors for *A. lumbricoides* infection. Interestingly, geophagia was not a risk factor for hookworm.

Saathoff *et al.* (2002) studied the association between geophagy and *A. lumbricoides* infection in rural schoolchildren from northern KwaZulu-Natal, South Africa. Prevalence of geophagy was found to be 39% in boys, and 53% in girls. The baseline prevalence of *A. lumbricoides* infection was greater in geophagous pupils (OR = 1.46; 95% CI 1.04–2.03).

**Bacterial infections**

**Leptospirosis**

Leptospirosis is a worldwide zoonosis with human contamination usually occurring after contact with water and/or soil that contains urine of infected rats and other animals (Ferguson, 1990; Ciceroni *et al*., 2000). The severity of this disease varies from mild to rapidly fatal, the latter being characterized by acute renal failure and haemorrhagic syndrome (Ferguson, 1990; Watt, 1997). In Western countries, leptospirosis is uncommon and mainly occurs in farmers and individuals who indulge in water-related activities. In tropical countries, leptospirosis can be up to 1000 times more frequent and risk factors may differ. Bovet *et al.* (1999) conducted a 1-year population-based case-control study to investigate the frequency and associated risk factors of leptospirosis in the entire population of the Seychelles (*n* = 74 331). This study was conducted between April 1995 and March 1996. All physicians within the Seychelles were informed, and requested to refer all patients suspected of having leptospirosis to one of two main hospitals in the country. Average rainfall over the 14 days preceding recruitment was obtained for all cases and controls from the National Meteorological Services. Blood samples were taken from cases within 24 h of admittance to hospital, with a further sample taken 2–4 weeks later. A single blood sample was taken from controls. Blood analysis was performed by the microagglutination test (MAT) (Perolat & Baranton, 1990; cited in Bovet *et al*., 1999).

The data were analysed by a multivariate logistic regression model that used both matched and un-matched case-control data. The most important risk factors included a skin wound within the last 14 days, and the drinking of home-brewed beer/alcohol. However, soil-related activities proved very strong risk factors. Using the matched data as an example, gardening within the last 14 days was shown to increase risk of leptospirosis by a factor of 10, although the confidence intervals were wide (OR = 10.5; 95% CI 2.46–44.7; *P* = 0.001). Soil that is often wet around the home increased the risk eight-fold (OR = 8.00; 95% CI 1.84–34.8; *P* = 0.006), but again the confidence intervals were wide. Activities in forests (OR = 4.50; 95% CI 0.97–20.8; *P* = 0.054) and walking barefoot near the home (OR = 3.00; 95% CI 1.09–8.25; *P* = 0.033) were also reasonably significant risk factors.

In a similar matched case-control study in Nakornratchasima Province, northeastern Thailand, Tanganakul *et al.* (2000) confirmed a range of agricultural and soil contact activities as risk factors for leptospirosis. The aim of this study was to identify potentially modifiable (i.e. risk management) risk factors and was carried out by the Thai Ministry of Public Health and the Royal Thai Army.

Nakornratchasima Province is a rural district, with the majority of the population engaged in farming. The study design was hospital-based, with controls matched to confirmed cases by gender, age (± 5 years), and neighbourhood of residence. Initially, analysis was performed by a matched-pairs univariate logistic regression model. A number of risk factors associated with rice cultivation were found to be significant including: (i) ploughing in a wet field (OR = 6.30; 95% CI 2.10–19.2); (ii) fertilizing a wet field (OR = 4.40; 95% CI 2.00–9.50); (iii) pulling out rice sprouts in a wet field (OR = 4.40; 95% CI 1.70–11.3); (iv) transplanting rice sprouts in a wet field (OR = 3.50; 95% CI 1.60–7.90); and (v) having cuts on feet whilst in paddy fields (OR = 2.60; 95% CI 1.10–6.20). However, because this is a water/soil environment it is difficult to put these increases in risk down to contact with soil *per se*.

**Q fever**

Q fever is a zoonosis caused by *Coxiella burnetii* and is endemic worldwide in a variety of birds, wild and domestic mammals, and arthropods (Maurin & Raoult, 1999). In most countries, humans are infected with *C. burnetii* by direct contact with aerosols generated during parturition of domestic ungulates. However, *C. burnetii* is highly resistant to desiccation in the environment, thus environmental exposure can occur (Gardon *et al*., 2001). This was investigated by a hospital-based case-control study in French Guiana conducted between January 1998 and June 2000 (Gardon *et al*., 2001). Controls were matched to new cases that presented themselves at the local clinics by age, sex, date of disease, and area of residence.
Blood samples were taken from all participants and assayed for the presence of C. burnetti antibodies, determined by complement fixation in cattle, dogs and cats (Dupuis et al., 1985; cited in Gardon et al. 2001). In total, 60 cases and 105 controls were selected for the study. A multivariate logistic regression analysis showed that working in the building trade increased risk of Q fever by three and a half times (OR = 3.54; 95% CI 1.10–11.0; P = 0.03). The inference can be made that these individuals are in close contact with soil. However, further work is needed to build on this evidence.

Legionnaires’ disease

Although transmission of Legionella bacteria has been associated primarily with airborne transmission from cooling towers, shower heads, and other aerosolizing devices (Fiore et al., 1998), there is increasing evidence that one species (Legionella longbeachae) is associated with gardening and the use of potting compost (Steele et al., 1990a,b; Koide et al., 1999; Duchin et al., 2000). Soil surveys conducted in Australia found that 73% (n = 33) of potting soil samples tested positive for Legionella longbeachae (Steele et al., 1990b). A survey of 17 potting composts in Japan returned positive results for Legionella longbeachae in eight samples (Koide et al., 1999). However, a similar survey in the UK (19 samples) yielded negative results.

At this stage it is difficult to say whether potting compost is a major reservoir for Legionella longbeachae because of the small sample sizes examined and the preliminary nature of the surveys conducted. Further evidence, including epidemiological evidence to determine if gardening with potting compost is a risk factor for Legionnaires’ disease, is required.

Mycoses

The impact of pathogenic fungi within the waste industry is reasonably well documented (Breum et al., 1996). However, the health impacts on people who come into close contact with soil (farmers, gardeners, construction workers) is less well characterized, or even underestimated (Spiewak, 1998). Several case reports have suggested an association between extensive contact with soil and development of fungal diseases (De Groote et al., 2000; Proctor et al., 2001; Komatsu et al., 2003). B. dermatitidis can cause a wide spectrum of clinical illness, ranging from asymptomatic infection to disseminated disease that involves the central nervous system, genitourinary tract, skin, liver, spleen, bone, heart, adrenal glands, and gastrointestinal tract (Abernathy, 1959; Sarosi et al., 1986; Meyer et al., 1993). Blastomycosis is contracted by inhalation of spores. De Groote et al. (2000) describe two case reports of blastomycosis, where the patients had been involved in a prairie dog relocation project. In the course of this work, they had both had extensive contact with contaminated soil. Proctor et al. (2001) performed a case-control risk analysis of a cluster of pulmonary blastomycosis cases in a rural community, Wisconsin, USA. Multiple earth-disturbing/digging activities were found to be statistically associated with illness (P < 0.01). However, this study only included eight cases and thus provides limited weight of evidence.

Fractures and height loss

Chronic low-level exposure to cadmium may promote calcium loss via urinary excretion, with bone disorders such as osteomalacia or osteoporosis being late manifestations of severe cadmium poisoning. In the general population, bone lesions have occurred only in Japan (Järup et al., 1998). However, in only one study was an attempt made to relate relatively small concentrations of cadmium in the soil to bone density and risk of fractures (Staessen et al., 1999). This analysis was based on data from a previous cross-sectional study that looked at
cadmium in Belgium (CadmiBel study; Staessen et al., 1991; cited in Staessen et al., 1999). In the follow-up study, 1014 individuals out of 1419 randomly selected participants agreed to take part. Measurements of bone density were taken, and participants also provided a 24-hour urine sample. The same standardized questionnaire that had been used in the CadmiBel study was used to collect updated information on participants’ lifestyle and medication intake. Soil samples were taken from 307 gardens owned by CadmiBel participants. The relative risks of fracture and height loss were calculated if the concentration of cadmium in the soil was doubled. For male participants, risk of fracture was found to be significant (OR = 1.39; 95% CI 1.04–1.86; P = 0.024), with risk of height loss elevated, but weakly significant (OR = 1.30; 95% CI 0.98–1.74; P = 0.072). For female participants, both risk of fractures (OR = 1.54; 95% CI 1.19–2.00; P < 0.001) and risk of height loss (OR = 1.29; 95% CI 1.02–1.64; P = 0.035) were significant. This study suggested that even at a relatively small degree of environmental exposure, cadmium can promote skeletal demineralization which leads to increased bone fragility and raised risk of fractures. However, it was assumed that the influence of age on bone density was minimal between the first study (1985–89) and the second study (1991–94). This aspect requires further investigation. It is also uncertain whether the soil samples taken suitably capture the exposure to Cd experienced by individuals.

Discussion and conclusions

This review has focussed on individual-level evidence for an association between exposure to soil and human health outcomes. Following the methodology of Goldbohm et al. (2006), weight of evidence for epidemiologic associations between soil and human health was greatest for specific cancers and enteric parasites (although E. coli was notable for its absence). As individual-level studies, the associations being investigated were often very specific and had been conceived from explicit prior evidence such as hospital records or case reports. Only one of the studies reviewed covered issues of relevance to the current worldwide situation (Kilian et al., 2001) where contaminated soils present a major issue in many urban areas. The majority of the studies had been undertaken in developing and/or rural subsistence communities where individuals are unlikely to have moved many times during their lives. This meant that geochemical or biological measurements taken around their homes were likely to have relevance to their current situation. This issue was highlighted by Montella et al. (1996), who showed no significant association between cancer and soil in the current area of residence in Italy, but the converse for soil in the place of birth. This is an important issue for aggregate-level studies that attempt to associate geochemical parameters to long-latency disease outcomes. Cancer epidemiology has shown early life, or cumulative exposure, to be the most appropriate predictors of cancer incidence (Fletcher et al., 2004b). Also, none of the studies featured actually attempted to measure soil exposure directly, i.e. the actual amount of soil (or soil contaminant) entering individuals’ bodies. In all cases, a proxy measure was taken by questionnaire—a tool very prone to bias. A questionnaire is often the only way to measure a large number of exposure scenarios, especially when data are collected retrospectively. However, it is good epidemiologic practice to take a current measurement (biomarker) of exposure as well as collect questionnaire data. Soil exposure can be measured for both contact and ingestion (e.g. McArthur, 1992; Calabrese et al., 1997; Lanphear et al., 1998; Pakkanen et al., 2003). Such measurements would provide some level of validation.

The findings of Bardos (2003), and the over-reliance on animal data for soil risk assessment both in the UK and the USA, seem to be upheld by the small number of studies identified in this review (only one study looked at contaminated land; Kilian et al., 2001). The majority of studies covered in this review relate to disease in developing countries, and outcomes associated with soil biology rather than chemistry (the main concern of contaminated land regulators, e.g. DEFRA & Environment Agency, 2002). Even within the studies identified, there is very little direct evidence for exposure to soil (or soil contaminants) as a risk factor for disease. At the same time, there is even less published evidence suggesting the contrary. Lack of evidence should not be thought of as evidence against an outcome-based approach. We have seen that individual-level epidemiologic data have been used successfully for risk management of a variety of endemic diseases, including goitre (El-Sayed et al., 1998), cyclosporiasis (Koumans et al., 1998; Bern et al., 1999), and leptoospirosis (Bovet et al., 1999; Tangkanakul et al., 2000). There is no reason why this approach would not aid risk management in contaminated land situations in the EU, USA and elsewhere.

Considering the economic burden of soil remediation, more investigations are needed into soil/soil-related activities as risk factors for disease outcomes associated with contaminated land, rather than an over-reliance on toxicological dose-response relationships and the uncertainties associated with intra- and interspecies extrapolation. It may not be possible to separate out health outcomes associated with specific soil contaminants. However, it should be possible to investigate the burden of disease at an individual level based on time spent on soil-related activities, or even measurements of soil ingestion/contact time. This step would seem a logical site-specific approach, once an inhabited site has been deemed ‘contaminated’ by toxicological models (Stephens et al., 2004; Hough et al., 2006). In this situation, the natural progression would often be to perform an ecological study. Individual-level studies are certainly a longer process than those at the aggregate-level, with even retrospective studies in this review taking 2 or 3 years, minimum, from instigation to publication. However, individual-level approaches are considered to be the “gold standard” where weight of evidence is concerned.
thus they could provide more appropriate data in the long term (Goldbohm et al., 2006).

In recent years, there has been a strong interest in ‘environment and health’ from a number of funding bodies. I am sure, however, that the majority of research proposed will have been aggregate-level studies that might or might not have been appropriate to either the exposure, or the outcome of interest. There is a need, highlighted by the significant gaps in knowledge outlined in this review, for individual-level study of specific health outcomes that might be influenced by soil factors (i.e. to come at the problem from the angle of the outcome, rather than the geochemistry). These data could be used to derive more appropriate dose-effect relationships, and hence reference doses, for risk assessment scenarios such as contaminated land.

References


© 2007 The Author


and chemical components in street-level and rooftop \( \text{PM}_1 \) particles in Helsinki. *Atmospheric Environment*, 37, 1673–1690.


