ORIGINAL ARTICLE

Hazard potential ranking of hazardous waste landfill sites and risk of congenital anomalies

M Vrijheid, H Dolk, B Armstrong, G Boschi, A Busby, T Jorgensen, P Pointer, and the EUROHAZCON collaborative group

Occup Environ Med 2002;**59**:768–776

See end of article for authors' affiliations

Correspondence to: Dr M Vrijheid, Unit of Radiation and Cancer, International Agency for Research on Cancer, 150 Cours Albert Thomas, 69372 Lyon Cedex 08, France; vrijheid@iarc.fr

Accepted 10 April 2002

Background: A 33% increase in the risk of congenital anomalies has been found among residents near hazardous waste landfill sites in a European collaborative study (EUROHAZCON).
Aims: To develop and evaluate an expert panel scoring method of the hazard potential of EUROHAZCON landfill sites, and to investigate whether sites classified as posing a greater potential hazard are those with a greater risk of congenital anomaly among nearby residents relative to more distant residents.
Methods: A total of 1270 cases of congenital anomaly and 2308 non-malformed control births were selected in 14 study areas around 20 landfill sites. An expert panel of four landfill specialists scored each site in three categories—overall, water, and air hazard—based on readily available, documented data on site characteristics. Tertiles of the average ranking scores defined low, medium, and high hazard sites. Calculation of odds ratios was based on distance of residence from the sites, comparing a 0–3 km "proximate" with a 3–7 km "distant" zone.
Results: Agreement between experts measured by intraclass correlation coefficients was 0.50, 0.44, and 0.20 for overall, water, and air hazard before a consensus meeting and 0.60, 0.56, and 0.53 respectively after this meeting. There was no evidence for a trend of increasing odds ratios with increasing overall hazard or air hazard. For non-chromosomal anomalies, odds ratios by water hazard category showed an increasing trend of borderline statistical significance (p = 0.06) from 0.79 in the low hazard category, 1.43 in the medium, to 1.60 in the high water hazard category.
Conclusions: There is little evidence for a relation between risk of congenital anomaly in proximate relative to distant zones and hazard potential of landfill sites as classified by the expert panel, but without external validation of the hazard potential scoring method interpretation is difficult. Potential misclassification of sites may have reduced our ability to detect any true dose–response effect.

The European collaborative study EUROHAZCON showed a 33% increase in risk of congenital malformation for residents living within 3 km of a hazardous waste landfill site,[1] pooling information from 21 sites. However, it is likely that sites differ in their hazard potential, because of a complex of site characteristics including age and size of the site, waste characteristics, geology, hydrogeology, climate, and engineering and management of the site.[2–4] If sites with higher hazard potential were found to be associated with a greater risk of congenital anomaly among nearby residents, this would potentially strengthen a causal interpretation of the overall association, and allow better appreciation of the risks that may be associated with different types of sites. In this paper we present the development and evaluation of an expert panel scoring method to score the relative hazard potential of landfill sites included in the EUROHAZCON study, based on readily available, documented data on site characteristics. We follow up the first EUROHAZCON findings by investigating whether sites classified as posing a greater potential hazard by the expert panel scoring method are those with a greater risk of congenital anomaly among nearby residents relative to more distant residents.

Previous large multisite studies in the USA have investigated risk of congenital malformation in relation to hazard categories of sites[5 6] or exposure indices incorporating hazard scores of sites.[7 8] In some of these studies higher risks have been found related to higher hazard sites,[5 6] or higher exposure indices,[7] adding support to evidence for possible causal relations. Several of these studies[5 7 8] have been able to use information from existing systematic scoring systems for hazard ranking of waste sites, developed as part of large site assessment programmes such as Superfund[9] or public health assessment programs.[10] Most existing US ranking systems were not suitable for use in our study because they required detailed information from site investigations. In Europe, no systematic and consistent site assessment procedure is in

Main messages

- The expert panel assessment of landfill sites presented in this study may, with some improvements, be a feasible method for assessment of hazard potential of landfill sites in future studies.
- There is little evidence for a relation between risk of congenital anomaly near landfill sites and hazard potential of the sites as classified by an expert panel.
- The main limitation of the expert panel hazard potential scoring method as presented in this paper is the absence of external validation of landfill exposure.

Policy implications

- In Europe, there are no systematic and consistent assessments of the hazard potential of landfill sites available for use in epidemiological surveillance studies or prioritisation of intervention, nor is information on waste inputs readily available.
- A causal relation between residence near hazardous waste landfill sites and risk of congenital anomalies does not receive any further support from this study.

Table 1 Landfill questionnaire response

Items included in questionnaire	Response*	%
Start and closure dates	19	95
Total site area	18	90
Total quantity of waste in place: volume or weight, depth	11	55
Hazardous waste quantity or % of total waste classified as "hazardous"	17	85
Types of hazardous and industrial waste deposited	20	100
Containment/lining	20	100
Covering	20	100
Capping	20	100
Leachate collection system	20	100
Leachate monitoring	20	100
Soil type and permeability	20	100
Groundwater depth	14	70
Groundwater monitoring	20	100
Groundwater contamination	11	55
Public drinking water supply extraction points within 3 km	18	90
Private water supply extraction points within 3 km	18	90
Surface water: type and distance	19	95
Surface water monitoring	20	100
Surface water contamination	9	45
Landfill gas control system	17	85
Landfill gas monitoring	17	85
Landfill gas migration	8	40
Complaints about smells and odours from the landfill	10	50
Rainfall	15	75
Land use for recreation and/or food consumption within 3 km	18	90

*Number and percentage of sites for which information on the questionnaire item was obtained.

place,[11] and information on waste inputs, particularly for the 70s and 80s, is very incomplete. We chose to develop an expert scoring method as the most feasible method for hazard potential assessment.

METHODS

Study design

The EUROHAZCON study is a multicentre case–control study which uses data from seven existing regional, population based congenital malformation registers in five European countries (Belgium, Denmark, France, Italy, UK). The methodology of the study and findings regarding risk of congenital anomaly in relation to distance of residence from sites are described in detail elsewhere for non-chromosomal[1] and chromosomal anomalies.[12]

In summary, we identified 21 landfill sites which contained "hazardous" waste of non-domestic origin (as defined in the EC directive 91/689 on Hazardous Waste[13]) in the regions covered by the participating centres. Each of the participating centres had found a collaborating landfill specialist in their region who could help them identify eligible sites and gather relevant information. Study areas were defined as 7 km zones around each site. If study areas of two or more sites were overlapping, study areas were combined to form one large study area. In this way, 15 study areas were defined around the 21 landfill sites, with three study areas containing more than one site. One of these study areas (area 14) was excluded because geographic site coordinates used in initial analyses proved incorrect, resulting in a total of 20 landfill sites in 14 study areas. (Exclusion of site 14 did not change findings published previously: the odds ratio for living within 3 km of a site including site 14 was 1.33 (95% CI 1.11 to 1.59) for non-chromosomal anomalies[1]; excluding site 14 this estimate was 1.34 (95% CI 1.12 to 1.60).) Within each study area, a 0–3 km "proximate" zone was defined on the advice of the collaborating specialists, to represent the zone of most likely exposure. In analyses this 3 km proximate zone was compared with a 3–7 km "distant" zone.

Cases included all malformed live births, stillbirths, and fetal deaths from 20 weeks gestation, and termination of pregnancy following prenatal diagnosis, born to mothers resident in one of the study areas, and born before 31 December 1994 and after five years of operation of the landfill site.[1] Cases with neoplasms, metabolic diseases, familial syndromes, minor malformations, and deformations were not included. Controls were non-malformed live births, approximately two per case, and selected from the same year of birth and same 7 km study area as the case.[1] They were either selected randomly from population registers, or by using one-to-one matching to facilitate the selection of controls in the absence of suitable population registers. Cases and controls were located geographically using addresses or postcodes at birth, with an accuracy of 100 m or better.

Landfill questionnaire

A questionnaire was completed for each of the study sites by the local waste authority responsible for the sites or their regulation. This questionnaire aimed to collect information that was readily available from existing documentation held by the waste site regulator, operator, inspector, and other relevant parties. Site visits were not carried out. Table 1 lists items included in the questionnaire.

The landfill questionnaire gave reasonably complete information on age and size of the EUROHAZCON study sites, soil type, and engineering and monitoring practices. Response rates for these items varied between 85% and 100% (table 1). Items related to total quantity of waste in place, contamination of ground or surface water, off site migration of landfill gas, and complaints about smells and odours were least well completed (40–55% response). For the majority of sites some monitoring results of either leachate, ground water, surface water, or landfill gas were available, but this type of information was not easily comparable between sites: monitoring was carried out for different substances, with different frequencies, on and off site, and in different years either during the study period or before. Summary reports of site investigations and monitoring were available for only six sites. Sites had all been reported to contain hazardous waste (as defined through the EC directive), but the amount of detail in the information deposited collected through the questionnaire on exact types and quantities of wastes was very variable (table 2). In most cases information on hazardous wastes deposited was limited to the types of industries from which the wastes originated.

Table 2 Site characteristics

Site	Operational years	Total site area	Types of industrial and/or "hazardous" waste
1	1962–72	1.1 ha	Tannery wastes with primarily ammonium and chromium compounds, tarry residues, oils, halogenated solvents
2	1950–74	2.2 ha	Chlorinated solvents, tar, phenols, cyanide, organic solvents
3	Pre1961–present	31 ha	Special and industrial wastes including oils, acids, alkalis, effluent/contaminated water and sludges, paint, leather phosphate, pesticide, electroplating wastes
4	1956–85	1.42 ha	Unspecified liquid and solid wastes
5	1979–88	10 ha	Various industrial toxic wastes including heavy metals, solvents, adhesives, varnishes, painting residues
6	1964–86	4.6 ha	Unspecified industrial wastes
7a	1955–60	10 ha	Industrial wastes: radium, residues from copper production
7b	Pre1970–?	2.1 ha	Copper
7c	1953–present	14.6 ha	Industrial wastes, chemical wastes from industrial water treatment
8	1966–present	100 ha	Industrial wastes: paint, rubber, ink, leather wastes, metal compounds, sludges
9	1978–84	2 ha	Unspecified hazardous industrial wastes
10	1974–83	7.6 ha	Incinerator waste, fly ash, and contaminated soil/sludges
11	Pre1950–85	1 ha	Heavy metals, wastes from production of sulphuric acid
12	1984–94	2.7 ha	Industrial, special and restricted waste
13a	1972–present	6.5 ha	Industrial waste: licensed to take inorganic and organic acids, alkalis, toxic metal compounds, miscellaneous chemical waste, treatment sludge, printing waste, tars, dyes, paints
13b	1973–89	12.5 ha	Licensed to take long list of special wastes including toxic metal compounds, cadmium, lead, mercury, adhesives, paint, tar, glue, printing waste, miscellaneous chemical waste
13c	1981–94	7.1 ha	Unspecified industrial and special wastes
13d	1983–93	10.1 ha	Industrial wastes including asbestos, boiler and flue cleaning, polymerisation products, adhesives, glue, and rubber wastes
15a	1955–present	29 ha	Various types of industrial waste: heavy metals, acid and alkali wastes, paint washings and solvents, industrial treatment sludges, tannery sludges, contaminated soils
15b	1935–68	2.5 ha	Waste from chemical works and chromium processing industry: chromium III and VI

Expert panel scoring

A panel of four was established from the group of collaborating regional landfill specialists on the basis of their varying geographic origin and interest in contributing expertise in hazard potential issues. Their collective expertise included fields of environmental chemistry, environmental and landfill engineering, hydrogeology, soil and ground water pollution, and risk assessment. This "expert panel" consisted of two landfill specialists who worked for regional environment agencies (in Scotland and Denmark) and two who worked for waste disposal companies (in England and Italy). The expert from Scotland had first hand knowledge for sites 15a and 15b and the expert from Denmark for sites 1 and 2. The other two experts had no first hand knowledge about any of the study sites, beyond that which they gained participating in the study. None of the experts were directly involved in the operation of any of the sites.

Results of the landfill questionnaire were summarised in a site description document and sent to the members of the expert panel. Each expert was asked to score each landfill site on the basis of the information provided in the site descriptions. Experts were blind to results of analyses of risk of congenital anomaly in relation to distance from each site. Sites were scored on a scale from 1 (low hazard) to 5 (high hazard) in three independent categories: water, air, and overall hazard table 3). The *water hazard scoring* aimed to reflect the ease with which hazardous materials can escape via the water route (groundwater and surface water), and the potential for the nearby population to come into contact with the water (via drinking water, surface water, recreation). The *air hazard scoring* aimed to reflect the ease with which hazardous substances in both vapour and particle form may be emitted into the air. The *overall hazard scoring* aimed to reflect a site's overall potential to cause exposure of nearby residents relative to other sites. A large, old, badly managed site with many reported problems, for example, would receive a higher overall score than a well managed, small site.

In a subsequent meeting the members of the expert panel discussed differences between their scores for each site and were given the chance to consult additional documentation on the sites such as inspection and monitoring reports and site maps. During the meeting, initial scores were changed when discussion between experts led to a consensus, when first hand knowledge from one of the experts changed the opinion of the others, or when the information given in the summary description proved to have been misinterpreted by one or more of the experts. As an example of the latter, one site was judged of low air hazard by one expert because a gas collection system was present, whereas the other experts had noted that the gas collection system was installed after the study period ended and scored the air hazard higher. The first expert increased his score at the meeting. As an example of first hand knowledge leading to a change in scores, a site for which groundwater pollution had been detected and which was near a private drinking water well was judged of high water hazard by three experts. First hand knowledge of the fourth expert clarified that the groundwater flow was away from the drinking water well and the others lowered their scores.

Final scoring and ranking

Final hazard scores (after changes were made) of the four experts were averaged to form the final overall, water, and air hazard scoring. In study areas containing more than one site, different hazard scores were given to different sites, which made the assignment of one score to the exposure zone in those study areas problematic. Study area matching of cases and controls meant that only one score could be assigned to each study area. Within study area classifications were not possible. It was decided that if 3 km "proximate" zones around sites did not overlap in these multiple site areas, the average hazard score of the sites, weighted by the proportion of controls in the proximate zone around each site, most accurately represented the hazard of the proximate zone in the study area. If the 3 km zones did overlap, the score of the highest scoring site was applied to the 3 km proximate zone. This algorithm was developed after consultation with members of the expert panel, but it was noted that it was not possible to be confident about how hazards from multiple sites would affect exposure of residents in an area.

High, medium, and low hazard categories were created using tertiles of the hazard scores as cut off points. This resulted in categories of five study areas each. After exclusion of site 14, low

Table 3 Expert panel scoring guide

Hazard in relation to water pathway:
Leaching and run-off of chemicals from the landfill site may cause ground and surface water contamination which would form a hazard for residents via drinking water consumption, other domestic water uses, and use of land in the vicinity of the site (i.e. food growing, recreation). The hazard scoring related to this water pathway reflects the ease with which hazardous materials can escape via the water route, and the potential for nearby population to come into contact with the water (via drinking water, surface water, recreation). Factors such as drinking water supply, land use, soil characteristics, distance to groundwater and surface water, and management practices such as leachate collection, lining of the site, monitoring and contamination of groundwater and surface water, are documented in the Site Description Document.

Water hazard scoring guide:

1	2	3	4	5
Low hazard		**Medium hazard**		**High hazard**
– Evidence of no contamination of ground or surface water, or – No water monitoring performed, but low potential hazard based on available information (site engineering, soil type, management of the site, size of the site, etc).		– Evidence for some off-site contamination of ground or surface water, and/or some potential for drinking water contamination, or – No, or limited water monitoring performed, but medium potential for off-site contamination on basis of other available information.		– Evidence of high off-site contamination of ground and/or surface water, and/or potential for drinking water contamination, or – No, or limited water monitoring performed, but high potential hazard relating to the water pathway based on available information.

Hazard in relation to air pathway:
Landfill sites may be a source of airborne toxic chemical contamination by evaporation or via windblown particles. The hazard scoring of the air pathway reflects the ease with which hazardous substances in both vapour and particle form may be emitted into the air. Migration of landfill gas is of importance because it may carry along waste vapours, such volatile organic substances. Factors such as the presence of a gas collection system, evidence of gas migration, and evidence of migration of other substances, are documented in the Site Description Document. Waste management practices such as capping and covering of the site are important in relation to the potential of dust and particles being blown off-site.

Air hazard scoring guide:

1	2	3	4	5
Low hazard		**Medium hazard**		**High hazard**
– Evidence of no off-site migration of landfill gas or air pollutants, or – No air monitoring performed, but low potential hazard based on information available (site engineering, quantities of waste, age of site, etc.).		– Evidence for some off-site migration of gas or air pollutants, or – No, or limited air monitoring, but medium potential for off-site contamination on basis of available information.		– Evidence of off-site migration of high levels of landfill gas and/or other air pollutants or, – No, or limited air monitoring performed, but high potential hazard based on other available information.

Overall hazard potential:
The scoring of the overall hazard of the sites reflects the overall potential of a site to cause exposure of nearby residential populations. A large, old, badly managed site with many reported problems would get a high overall score and a well-managed, small site with no reported problems would be assigned a low overall hazard score. Factors such as age of the site, size of the site, quantities of waste present, actions taken to prevent leachate and landfill gas emissions, adequacy of the monitoring at the site, and contamination problems, are documented in the Site Description Document.

Overall hazard scoring guide:

1	2	3	4	5
Low hazard		**Medium hazard**		**High hazard**
– Relatively new well-managed site with few reported problems. – Relatively small quantity of waste in place. – Adequate measures to prevent gas and leachate migration in place since start of site. – Adequate monitoring of the site.		– Site of medium size and age. – Some reported problems. – Adequate measure to prevent gas and leachate migration have been taken over time/some measures have been taken but not adequate. – Some monitoring, but not adequate.		– Site with large quantities of waste and/or old, uncontrolled site, and: – No measures taken to prevent off-site migration of landfill gas or leachate. – No routine monitoring. – Many problems reported.

hazard categories for overall, water, and air hazard contained four study areas each.

Analysis

In order to assess the agreement between experts in both initial and final scores, intraclass correlation coefficients (ICC) were calculated by analysis of variance. In addition, the reliability of the average expert scores (ICC_k) was calculated. ICC and ICC_k are calculated as follows[14]:

ICC = variance between sites/(variance between sites + variance within sites)

= inter-rater agreement = reliability of single rater

ICC_k = variance between sites/((variance between sites + variance within sites)/k)

= reliability of mean of k raters = reliability of the average score.

An intraclass correlation coefficient (ICC) of 1 reflects perfect agreement between experts.

In order to investigate whether the hazard potential of a site modified the odds ratio for residence within 3 km from a site, odds ratios for living within 3 km from a waste compared to further away (3–7 km) were calculated in each of the three hazard categories (high, medium, low). All odds ratios were stratified by study area and year of birth and adjusted for maternal age and socioeconomic status using logistic regression models.[1] The likelihood test for the interaction term between hazard category as a numerical variable (1=low, 2=medium, 3=high hazard) and distance zone (0–3 versus 3–7 km) was then used to test for the statistical significance of the trend in odds ratio from low to high hazard category. In addition, the interaction term between continuous hazard score (for 14 study areas) and distance zone (0–3 versus 3–7 km) was used to test for linear trend in odds ratio with continuous hazard score.

Hazard scoring analyses were carried out for all non-chromosomal anomalies combined, all chromosomal anomalies combined, and the three malformation subgroups which showed statistically significant increased risks related to residence within 3 km from landfill sites in our previous work (neural tube defects, cardiac septal defects, and malformations of the great arteries and veins).[1]

RESULTS

Expert scoring

Initial scores

Few sites were given the score of 1 (low hazard)—the only ones were three sites which scored 1 for air hazard. Air hazard was generally scored lower than water hazard. Agreement between experts as measured by the intraclass correlation coefficient, was better for overall (ICC = 0.50) and water hazard scores (ICC = 0.44) than for air hazard (ICC = 0.20). The

Table 4 Final expert panel hazard scores—individual scores, average scores, hazard categories, and agreement between experts

		Overall						Water						Air						
		Expert						Expert						Expert						
Study area	Site	1	2	3	4	Hazard score (ave.)	Hazard category	1	2	3	4	Hazard score (ave.)	Hazard category	1	2	3	4	Min–max	Hazard score (ave.)	Hazard category
1	1	**3**	4	2.5	4	3.38	Medium	**3**	**4**	2.5	4	3.38	Low	**3**	**3**	1	3	1–3	2.50	Low
2	2	**4**	**4**	3.5	3	3.63	Medium	5	5	4	4	4.50	High	4	5	**3**	3	3–5	3.75	High
3	3	2	3	3	3	2.75	Low	3	3	3	3	3.00	Low	**3**	3	**3**	3	3–3	3.00	Medium
4	4	5	4.5	5	4	4.63	High	5	5	5	4	4.75	High	4	4	4.5	**4**	4–4.5	4.13	High
5	5	4	4.5	5	4	4.38	High	3	4	5	4	4.00	Medium	4	5	5	4	4–5	4.50	High
6	6	4	4	4.5	4	4.13	High	4	4	4.5	4	4.13	High	4	3	4.5	3	3–4.5	3.63	High
7	7a	**3**	3	4.5	3	3.38		4	4	4.5	3	3.88		**3**	2	3.5	3	2–3.5	2.88	
	7b	3	2	2	3	2.50		4	3	1.5	3	2.88		3	2	1	3	1–3	2.25	
	7c	3	3	3	4	3.25		4	4	3	4	3.75		3	**3**	3	3	3–3	3.00	
						3.38*	Medium					3.88*	Medium						3.00*	Medium
8	8	4	4	4	4	4.00	High	5	4.5	4	4	4.38	High	4	4	**4**	**4**	4–4	4.00	High
9	9	3	3	3	**3**	3.00	Low	3	4	3	4	3.50	Medium	2	2	3	3	2–3	2.50	Low
10	10	3	3	3.5	3	3.13	Low	3	**3**	3.5	3	3.13	Low	3	2	3	3	2–3	2.75	Low
11	11	4	3	3	4	3.50	Medium	3	3	3	4	3.25	Low	4	2	2	3	2–4	2.75	Low
12	12	4	3	3.5	3	3.38	Medium	4	**3**	4	3	3.50	Medium	4	3	**3**	**3**	3–4	3.25	Medium
13	13a	**3**	3	3.5	3	3.13		**3**	2	**3**	**3**	2.75		3	3	2.5	3	2.5–3	2.88	
	13b	3	4	3.5	3	3.38		4	4.5	4	**4**	4.13		3	3	3	3	3–3	3.00	
	13c	2	2	3	3	2.50		2	1.5	3	3	2.38		2	2.5	3	3	2–3	2.63	
	13d	2	3	3	3	2.75		2	2	3	3	2.50		**3**	3	3	4	3–4	3.25	
						3.29*	Low					3.65*	Medium						3.09*	Medium
15	15a	4	5	4.5	4	4.38		4	5	4.5	4	4.38		**4.5**	5	4.5	4	4–5	4.50	
	15b	4	4	4	4	4.00		4	5	4	4	4.25		**2.5**	**2.5**	**2.5**	**2.5**	2.5–2.5	2.50	
						4.13*	High					4.30*	High						3.21*	Medium
ICC†		0.60						0.56						0.53						
ICC_k		0.86						0.83						0.82						

*In study areas containing more than one site composite scores were calculated for the entire exposure zone within these study areas.
†Intraclass correlation coefficient.
Notes: scores of 1.5, 2.5, 3.5, 4.5 reflect where experts gave in-between scores: 1–2, 2–3, 3–4, 4–5.
Bold scores are those that have been changed during the expert panel meeting.

differences between the lowest and the highest expert score (measured on scale 1 to 5) given to a site also reflect this. For the majority of sites the difference between expert scores is one point or less in the overall (16 sites) and water scoring (12 sites), whereas in the air scoring only seven sites show one point or less difference between experts. Three sites show a difference of three points or more in the air hazard scoring.

Final scores

Table 4 shows the final hazard scores assigned by each expert to the study sites, as well as the average scores and the hazard category (low, medium, high) of each site. Scores that were changed during the expert panel meeting are emboldened in table 4. Few scores were changed in the overall and water hazard scoring: six and eight respectively. Differences between experts were greater for the initial air hazard scoring and 18 air scores for 11 sites were changed. As expected, changes made at the expert panel meeting improved agreement between experts. The intraclass correlation coefficient (ICC) for overall hazard scores increased from 0.50 to 0.60, water hazard scores from 0.44 to 0.56, and air hazard scores from 0.20 to 0.53. The number of sites differing by 1 point or less is 18 in the final overall hazard scoring, 15 in the final water hazard scoring, and 14 in the final air hazard scoring. Differences of two or more points are found for site 5 and 7b in the water score, and sites 1, 2, 7b, and 11 in the air score. The difference between the lowest and highest scoring expert was never more than 2.5 points in the final scores. The reliability of the average of the final scores of the four experts was high for overall ($ICC_k = 0.86$), water ($ICC_k = 0.83$), and air hazard scores ($ICC_k = 0.82$). Average scores covered a limited range with overall scores ranging from 2.50 to 4.63, water scores from 2.50 to 4.75, and air scores from 2.25 to 4.50 (table 3). The average final overall and water scores were highly correlated, with a correlation coefficient of 0.86. Correlations between overall and air (0.76) and water and air (0.62) were not as strong. All correlation coefficients were statistically significant ($p < 0.01$).

Relation between risk of congenital anomaly and hazard score

In table 5 the odds ratios (ORs) for living within 3 km from a landfill site compared to living further away from a site are presented pooled for all study areas and by low, medium, and high hazard categories. There was no evidence for a trend of increasing odds ratios with increasing overall hazard or air hazard. Odds ratios by water hazard category show an increasing trend of borderline statistical significance ($p = 0.06$) from 0.79 (0.51–1.21) in the low hazard category, 1.43 (1.10–1.86) in the medium, to 1.60 (1.16–2.21) in the high water hazard category. Testing the linear trend in the odds ratios for the 14 study areas with continuous hazard scores gave broadly similar results, although water hazard now suggests only a very weak and not statistically significant increasing trend ($p > 0.2$).

Odds ratios for chromosomal anomalies showed a similar pattern over the various hazard categories to those for non-chromosomal anomalies (table 6). Again only water hazard showed some weak, and not statistically significant, suggestion of an increase in odds ratios with hazard category.

In analyses of neural tube defects, cardiac septal defects, and malformations of the great arteries and veins, numbers of cases in different hazard categories were often small and confidence intervals wide, giving very limited power to test for differences between odds ratios (table 7). For neural tube defects odds ratios increased with air hazard category from 0.46 (95% CI 0.10 to 2.09) for low hazard, 1.93 (95% CI 1.23 to 3.02) for medium hazard, to 3.81 (95% CI 1.01 to 14.43) for high hazard, but this trend did not reach statistical significance ($p = 0.06$ for trend in three ORs). Odds ratios for

Table 5 Odds ratios for living within 3 km from a waste site by low, medium, and high hazard category—non-chromosomal anomalies

Hazard category	Study area	Distance zone	Cases	Controls	OR*	95% CI
All study areas		≤3 km	294	507	1.34	1.12 to 1.60
Undivided by hazard category		3–7 km	772	1801		
Overall hazard						
Low	3, 9, 10, 13	≤3 km	127	202	1.43	1.09 to 1.88
		3–7 km	400	867		
Medium	1, 2, 7, 11, 12	≤3 km	80	147	1.15	0.78 to 1.69
		3–7 km	167	400		
High	4, 5, 6, 8, 15	≤3 km	87	158	1.48	1.07 to 2.04
		3–7 km	205	534		
Trend in 3 ORs with hazard category					p=0.87	
Trend in 14 ORs with continuous hazard score					p=0.90	
Water hazard						
Low	1, 3, 10, 11	≤3 km	77	135	0.79	0.51 to 1.21
		3–7 km	158	224		
Medium	5, 7, 9, 12, 13	≤3 km	123	203	1.43	1.10 to 1.86
		3–7 km	423	1064		
High	2, 4, 6, 8, 15	≤3 km	94	169	1.60	1.16 to 2.21
		3–7 km	191	513		
Trend in 3 ORs with hazard category					p=0.06	
Trend in 14 ORs with continuous hazard score					p=0.23	
Air hazard						
Low	1, 9, 10, 11	≤3 km	73	91	0.95	0.62 to 1.48
		3–7 km	172	211		
Medium	3, 7, 12, 13, 15	≤3 km	182	337	1.48	1.19 to 1.85
		3–7 km	487	1317		
High	2, 4, 5, 6, 8	≤3 km	39	79	1.23	0.75 to 2.02
		3–7 km	113	273		
Trend in 3 ORs with hazard category					p=0.55	
Trend in 14 ORs with continuous hazard score					p=0.65	(negative trend)

*Adjusted for socioeconomic status and maternal age.

Table 6 Odds ratios for living within 3 km from a waste site by low, medium, and high hazard category—chromosomal anomalies

	Overall			Water			Air		
Hazard category	n*	OR†	95% CI	n*	OR†	95% CI	n*	OR†	95% CI
Low	105	1.53	0.88 to 2.68	46	0.98	0.34 to 2.80	41	1.12	0.42 to 3.01
Medium	51	1.28	0.56 to 2.91	108	1.55	0.91 to 2.63	139	1.51	0.97 to 2.34
High	48	1.65	0.83 to 3.29	50	1.66	0.85 to 3.23	24	1.15	0.36 to 3.61
Trend in 3 ORs with hazard category:			p=0.87			p=0.22			p=0.85
Trend in 14 ORs with hazard score:			p=0.62			p=0.37			p=0.93

*Number of cases.
†Adjusted for socioeconomic status and maternal age.

malformations of cardiac septa increased with water hazard (low hazard OR: 0.99, 95% CI 0.50 to 1.99; medium hazard OR: 1.57, 95% CI 1.02 to 2.42; high hazard OR: 2.02, 95% CI 1.07 to 3.83), but again this trend did not reach statistical significance (p =0.16 for trend in three Ors).

DISCUSSION

Evaluation of scoring method

The interpretation of our results concerning the presence or absence of a relation between risk of congenital anomaly and hazard potential is dependent on an evaluation of the validity of the hazard scoring system. The hazard potential of a landfill site is dependent on many factors, and importantly also their interrelation. In the context of this study we believe it inappropriate to categorise sites by individual site characteristics (for example, age, size, waste type, containment, or engineering method). Analysis of a large number of such individual characteristics would lead to interpretational problems related to multiple statistical testing and small numbers, particularly in the absence of strong independent evidence of the degree to which these characteristics each individually determine hazard potential. Instead we aimed to capture hazard potential by combining information on many characteristics in a single scoring method.

Little is known in the published literature about the validity of existing hazard potential ranking systems, even of well

Table 7 Odds ratios for living within 3 km from a site by low, medium, and high hazard category—malformation subgroups

	Overall			Water			Air		
Hazard category	n	OR*	95% CI	n	OR*	95% CI	n	OR*	95% CI
Neural tube defects									
Low	74	2.14	1.25 to 3.69	19	1.88	0.66 to 5.35	12	0.46	0.10 to 2.09
Medium	24	1.05	0.38 to 2.85	78	1.74	1.01 to 3.00	103	1.93	1.23 to 3.02
High	28	1.89	0.84 to 4.29	29	1.94	0.87 to 4.32	11	3.81	1.01 to 14.43
Trend in 3 ORs with hazard category			p=0.64†			p=0.95			p=0.06
Trend in 14 ORs with hazard score			p=0.74†			p=0.81			p=0.20
Cardiac septal defects									
Low	119	2.08	1.32 to 3.27	61	0.99	0.50 to 1.99	74	1.71	0.92 to 3.16
Medium	74	0.84	0.45 to 1.55	134	1.57	1.02 to 2.42	137	1.42	0.93 to 2.15
High	49	1.65	0.86 to 3.16	47	2.02	1.07 to 3.83	31	1.60	0.66 to 3.84
Trend in 3 ORs with hazard category			p=0.31†			p=0.16			p=0.68†
Trend in 14 ORs with hazard score			p=0.67†			p=0.32			p=0.31†
Malformations of the great arteries and veins									
Low	27	1.69	0.69 to 4.14	14	1.36	0.38 to 4.82	12	1.17	0.27 to 5.12
Medium	10	1.51	0.33 to 6.96	25	1.63	0.63 to 4.18	42	2.15	1.11 to 4.15
High	24	2.19	0.93 to 5.17	22	2.59	1.06 to 6.32	7	1.20	0.18 to 8.08
Trend in 3 ORs with hazard category			p=0.70			p=0.47			p=0.99
Trend in 14 ORs with hazard score			p=0.84			p=0.66			p=0.59

*Odds ratios stratified for study area and year of birth, not adjusted for socioeconomic status or maternal age.
†Negative trend.

used systems such as the USEPA Hazard Ranking System.[9] A method for external validation of our expert panel assessment was not available. It was not feasible to take measurements of individual chemicals in air, water, or soil near the study sites and there was no reliable documented information on such measurements. In the absence of reliable and feasible methods for determining exposure to landfills, the expert panel scoring of hazard potential of landfill sites, even though crude, was the best proxy available. We compared the expert panel scoring to an adaptation of one published ranking system which was developed for use with existing site documentation and did not require information from site visits.[15] Where there were differences, reasons could usually be found that pointed out deficiencies in the published ranking system. In particular, expert judgements were more able than the published ranking system to take into account the interrelations between factors. For example, the presence of a gas collection system was particularly important if significant biodegradable waste was present.

Expert panels have not been commonly used to assess environmental hazards, but they have proven useful in occupational settings to estimate exposures from job descriptions and titles where direct exposure measurements were not available.[16–19] The agreement between experts on a panel can give some indication of the reliability and therefore validity of a scoring method. Also, the more experts on a panel, the more repeatable, and therefore reliable, an average score will be (average score reliability for final scores was between 0.82 and 0.86 for our four member panel).[20 21] Agreement between experts in this study, measured by the interclass correlation coefficient, ranged from 0.20 (for air hazard) to 0.50 (for overall hazard) in the initial hazard ranking, increasing to 0.53 (for air hazard) to 0.60 (for overall hazard) after they had a chance to meet and discuss. Values of interrater agreement (reliability) between 0.40 and 0.75 have been reported as fair to good, values above 0.75 as excellent, and values below 0.40 as poor.[22] Interrater agreements reported in occupational studies rarely exceed the value of 0.7.[22] The final agreement found in this study falls within the range of interrater agreements reported, for example, in studies of pesticide applicants (0.4–0.8[19]), exposures of sawmill workers (0.40–0.68[18]), workers in various manufacturing industries (0.5–0.7[16]), and is higher than found in expert panels assessing metal exposures (0.2–0.5[17]) and various occupational chemical exposures (0–0.6[22]). Comparisons are problematic of course, since different methods for expert assessment have been applied in these different studies.

Landfill questionnaires gave reasonably complete information on site characteristics such as size, age, engineering, and management practices, but there was little documented data on actual waste types deposited, chemicals present in the site, and off site migration of substances from the sites. Information on types of waste present (that is, chemical composition) would probably have been of limited use to differentiate sites, even if available. The vast majority of sites took a mixture of chemicals and our ability to judge the relative teratogenic potential of different waste types is very limited. There are no strong prior hypotheses about which specific chemicals or chemical mixtures may cause congenital malformation, although many chemicals commonly present in landfills (organic solvents, heavy metals, pesticides) have shown teratogenic potential.[23] Also, teratogenic potential depends on dose and there exists insufficient information on this. Moreover, the composition of wastes entering a site may bear very little resemblance to that of trace contaminants present in leachate and landfill gas emissions from sites. Indeed, in order to judge a site's potential to generate landfill gas an estimate of the amount or proportion of biodegradable waste present in each site would have been of more use than a more detailed breakdown of waste types. First hand knowledge of sites was considered quite valuable by our expert panel. For example, additional knowledge of the direction of groundwater flow was used to judge potential risk to drinking water wells. First hand knowledge therefore addressed gaps in the questionnaire. First hand knowledge was only available for four sites and there were relatively more changes in scores for these sites (13 of 48 scores) than for other sites (19 of 192 scores). Ideally in future studies, site visits by one or more of the panel experts may give a better idea of the management of sites and adequacy of some pollution prevention measures, but cost effectiveness of such visits would need further evaluation.

Incomplete information may have resulted in either over or underestimation of the true relative hazard of sites. Where information that should normally be available is missing, this may indicate poorly managed sites with less pollution controls and therefore higher hazard potential. A study of the US EPA Hazard Ranking System on the other hand showed that missing information usually led to underestimation of hazard potential.[24] In our study, most data items were well completed, and the main issue was the limited scope of information available.

It was difficult to classify the hazard potential of study areas containing multiple sites with differing hazard potential scores. Experts agreed on an algorithm to classify these sites but the algorithm could not be validated. In future studies, dispersion modelling using meteorological, topographical, and hydrogeological information may be valuable in mapping patterns of relative exposure around landfill sites and could underpin the hazard potential assessments in multiple site areas as well as the definition of distance based exposure zones.

Interpretation of results regarding risk of congenital anomaly

Previous EUROHAZCON findings have shown an increased risk of congenital anomaly for residents living close to (within 3 km of) a hazard waste landfill site.[1 12] Potential sources of bias in the relation between distance of residence from sites and risk of congenital anomaly, including misclassification of exposure, ascertainment bias, migration bias, occupational and industrial exposures, and socioeconomic confounding, are discussed in our previous paper in detail.[1] The current findings show little evidence for relative risk of congenital anomaly close to landfill sites to be associated with the estimated hazard potential of landfill sites. Data show some evidence, although not statistically significant, of a trend of increasing relative risk of congenital anomaly with increasing water hazard of sites. The relation with water hazard could be a chance finding, as indicated by its low statistical significance. It provides suggestive evidence, however, that water is a more important exposure pathway than air for sites in this study, or that water is an equally or less important pathway than air, but easier to measure. There may be some reason to believe that water hazard was easier to classify than air hazard from the information available to the experts since agreement between experts on the initial water hazard scoring was considerably better than the initial air hazard scoring. Knowledge about pathways of potential exposure to landfill sites is as yet severely limited, adding to difficulties in interpretation of these findings.

Malformation subgroups analysed in relation to the hazard potential classification showed different patterns of risk with hazard potential: neural tube defects showed some evidence of a trend with air but not with water hazard, cardiac septal defects showed some evidence of a trend with water but not with air hazard. Although these findings may be caused by chance (again the trends reported were not statistically significant), they may alternatively indicate risks of different malformations occurring through different possible exposure pathways, possibly through exposures to different substances. This can only be resolved in a larger study with more detailed exposure assessment.

If misclassification of the relative hazard potential of one or a few sites occurred, this could have had an important impact on results regarding the risk of congenital anomaly risk near sites in each hazard category, especially if multiple site areas (for example, areas 13 and 15) and sites in the more densely populated study areas were misclassified. Such misclassification would reduce our ability to detect any true relation between risk of congenital anomaly and hazard potential.

Conclusion

We have shown the development of an expert panel hazard potential scoring method for an environmental exposure, and indicated ways in which the method could be improved in future studies of environmental exposures in general and landfill in particular. It is recognised that the hazard potential assessment presented in this paper has many limitations, the main one being the absence of an external validation method. However, the assessment method presented forms a basis for further developments and indeed expert assessment may be the only feasible way to assess the potential hazard of landfill sites in epidemiologic surveillance based studies in Europe.

Using the expert assessment method, we find little evidence for a relation between risk of congenital anomaly among proximate relative to distant residents and hazard potential of sites. This finding does not add support to a causal interpretation of the relation between distance from a waste site and risk of congenital anomaly. In the absence of external validation of the hazard potential scoring method, interpretation should be cautious. The extent of misclassification of hazard potential of sites is difficult to estimate and such misclassification may have reduced our ability to detect any true dose–response effect.

ACKNOWLEDGEMENTS

The work for this paper was carried out under a Research Fellowship for Martine Vrijheid from the Colt Foundation. Coordination of the EUROHAZCON study was funded by the European Commission DGXII BIOMED programme Concerted Action Contract BHM-94-1099. We thank Calum MacDonald (SEPA) for his work as a member of the expert panel. We are grateful to landfill specialists in participating regions (Michael Fogh, Tom McDonald, Isabel Melkebe, Marco Pellegrini, Fabio Del Soldato, Branko Druzina, and staff at Réseau Santé-Déchet) for their help in completing landfill questionnaires.

Authors' affiliations

M Vrijheid, H Dolk, B Armstrong, A Busby, Environmental Epidemiology Unit, Department of Public Health and Policy, London School of Hygiene and Tropical Medicine, London, UK
G Boschi, APS, Azienda Padova Sevizi, Spa, Padova, Italy
T Jorgensen, COWI Consulting Engineers and Planners, Odense, Denmark
P Pointer, Shanks & McEwan (Southern) Ltd, Milton Keynes, UK

EUROHAZCON Collaborative Group: Lenore Abramsky, North Thames (West) Congenital Malformation Register; Fabrizio Bianchi, Tuscany EUROCAT Register; Ester Garne, Funen County EUROCAT Register; Vera Nelen, Antwerp EUROCAT Register; Elisabeth Robert, France Central East Congenital Malformation Register; John ES Scott, Northern Region Fetal Abnormality Survey; David Stone, Glasgow EUROCAT Register; Romano Tenconi, North-East Italy EUROCAT Register

REFERENCES

1 **Dolk H**, Vrijheid M, Armstrong B, *et al.* Risk of congenital anomalies near hazardous-waste landfill sites in Europe: the EUROHAZCON study. *Lancet* 1998;**352**:423–7.
2 **Westlake K**. *Landfill waste pollution and control.* Chichester: Albion Publishing, 1995.
3 **Lisk D**. Environmental effects of landfills. *Sci Total Environ* 1991;**100**:415–68.
4 **Campbell DJV**. Environmental management of landfill sites. *Journal of the Institute of Water Environment Management* 1993;**7**:170–3.
5 **Croen LA**, Shaw GM, Sanbonmatsu L, *et al.* Maternal residential proximity to hazardous waste sites and risk of selected congenital malformations. *Epidemiology* 1997;**8**:347–54.
6 **Shaw GM**, Schulman J, Frisch JD, *et al.* Congenital malformations and birthweight in areas with potential environmental contamination. *Arch Environ Health* 1992;**47**:147–54.
7 **Geschwind SA**, Stolwijk JAJ, Bracken M, *et al.* Risk of congenital malformations associated with proximity to hazardous waste sites. *Am J Epidemiol* 1992;**135**:1197–207.
8 **Marshall E**, Gensburg L, Geary N, *et al.* Maternal residential exposure to hazardous waste sites and risk of central nervous system and musculoskeletal birth defects. *Epidemiology* 1997;**6**:S63.
9 **United States Environmental Protection Agency**. Hazard Ranking System—Final Rule. 40 CFR Part 300. *Federal Register* Friday, 14 December 1990, 51532-51667.
10 **New York State Department of Health**. *Human exposure potential ranking model for hazardous waste sites. Methodology report.* Albany, NY: New York State Department of Health, 1986.
11 **Department of the Environment**. *Health effects from hazardous waste landfill sites.* London: Department of the Environment, 1994.

12 **Vrijheid M**, Dolk H, Armstrong B, *et al*. Risk of chromosomal congenital anomalies in relation to residence near hazardous waste landfill sites in Europe. *Lancet* 2002;**359**:320–2.
13 **European Communities Council**. Council Directive of 12 December 1991 on Hazardous Waste (91/689/EC) OJ No L377/20. *Official Journal of the European Communities* 1991;L377/20, 31.12.91.
14 **Winer BJ**. Chapter 4. In: *Statistical principles in experimental design*, 2nd edn. New York: McGraw-Hill, 1971:261–96.
15 **JRB Associates**. *Methodology for rating the risk potential for hazardous waste disposal sites*. McLean, VA: JRB Associates, 1982.
16 **Goldberg MS**, Siemiatycki J, Gerin M. Inter-rater agreement in assessing occupational exposure in a case-control study. *Br J Ind Med* 1986;**43**:667–76.
17 **Rybicki BA**, Peterson EL, Johnson CC, *et al*. Intra- and inter-rater agreement in the assessment of occupational exposure to metals. *Int J Epidemiol* 1998;**27**:269–73.
18 **Teschke K**, Hertzman C, Dimich-Ward H, *et al*. A comparison of exposure estimates by worker rater and industrial hygienists. *Scand J Work Environ Health* 1989;**15**:424–9.
19 **de Cock J**, Kromhout H, Heederisk D, *et al*. Subjective assessment of pesticide exposure in fruit growing by experts. *Scand J Work Environ Health* 1996;**22**:425–32.
20 **Fleiss JL**. The measurement of inter-rater agreement. In: Fleiss JL, ed. *Statistical methods for rates and proportions*, 2nd edn. New York: John Wiley & Sons, 1981:212–36.
21 **Streiner DL**, Norman GR. *Health measurement scales. A practical guide to their development and use*. Oxford: Oxford Medical Publications, 1989.
22 **Benke G**, Sim M, Forbes A, *et al*. Retrospective assessment of occupational exposure to chemicals in community-based studies: validity and repeatability of industrial hygiene panel ratings. *Int J Epidemiol* 1997;**26**:635–42.
23 **Schardein JL**. *Chemically induced birth defects*, 3rd edn. New York: Dekker, 2000.
24 **Haness S**, Warwick J. Evaluating the hazard ranking system. *Journal of Environmental Management* 1991;**32**:165–76.

Hazard potential ranking of hazardous waste landfill sites and risk of congenital anomalies

M Vrijheid, H Dolk, B Armstrong, et al.

Occup Environ Med 2002 59: 768-776
doi: 10.1136/oem.59.11.768