Bull. Environ. Contam. Toxicol. (2004) 72:1182–1186

DOI: 10.1007/s00128-004-0368-2

Bulletin of Environmental Contamination and Toxicology

Finding Hazardous Waste Sites: Evaluation Using the Capture–Recapture Method

J. H. Lange,[1] Y.-F. Chang,[2] R. E. LaPorte,[3] E. O. Talbott,[3] M. Wang,[3] G. Mastrangelo[4]

[1] Envirosafe Training and Consultants, Post Office Box 114022, Pittsburgh, PA 15239, USA
[2] Department of Neurosurgery, UPMC Presbyterian, Suite B-400, 200 Lothrop Street, Pittsburgh, PA 15213, USA
[3] Graduate School of Public Health, University of Pittsburgh, 130 DeSoto Street, Pittsburgh, PA 15261, USA
[4] Department of Environmental Medicine and Public Health, Section of Occupational Health, University of Padova, Via Giustiniani, 2-35128 Padova, Italy

Received: 26 September 2003/Accepted: 10 March 2004

Hazardous waste sites (HWS) are considered an important environmental, occupational and public health hazard in the United States (Roney et al., 1999) and throughout the world (Saad, 2003). HWS in the United States have been traditionally defined by the federal government through the Resource Conservation Recovery Act and the Comprehensive Environmental Compensation and Liability Act with these definitions extended to state governments. Both federal and state governments locate and identify HWS. However, as with any form of counting there will be an underestimate due to the fact that all sites will not be located (Lange et al. 2003a). This leads to the question as to where we should look for these missed sites.

Using lists generated by federal and state government agencies the capture-recapture method (CRM) can be used to ascertain the number of HWS. This methodology has been used successfully in counting more difficult populations (wildlife) as well as in epidemiology (LaPorte et al.,1992; Lange and LaPorte, 2003). It has been recently suggested that this method can also be used for environmental and occupational health as well (Lange et al., 2003).

This investigation used the CRM to estimate the number of HWS for each county in Maryland and determine if income, population and/or manufacturing income are related to location of hazardous waste sites.

MATERIALS AND METHODS

These data were evaluated with population, manufacturing shipment costs and income. HWS lists (sources) were obtained from the federal Comprehensive Environmental Response Compensation Liability Information System-CERCLIS list (Environmental Protection Agency, 2003) and the state list for Maryland (Maryland Department of the Environment, 2002).

These sites were matched by site name and city/county location. Federal and state sites were matched to the time period of July, 2002. A two-source CRM was used to ascertaining an estimate of the true number of

Correspondence to: J. H. Lange

sites (ascertainment and undercount – missed sites) and confidence intervals (CI) at 95% (Lange et al., 2003a; McCarty et al., 1993). The number of missed HWS was listed by federal and state categorization for each county/location. Ascertainment represents the number of sites that are countable (estimate of the true number of sites). Income (per person), population (person per square mile – sqm) and manufacturing shipment costs were obtained from the US Census Bureau (2003).

Since Howard county had a negative number for missed state HWS and St. Mary and Somerset counties did not have manufacturing shipment cost reported, they were excluded in statistical analysis for some evaluations.

Spearman rank (SAS Institute, Inc., Cary, NC) was used for correlation analysis. Population, manufacturing shipments and income were normally distributed. State and federal undercounts and ascertainment could not be fitted to a normal distribution by Box-Cox transformation or other transformations (e.g. arc sine) (Sokal and Rohlf, 1995; SAS Institute, Inc., Cary, NC). Distribution was determined using the Shapiro-Wilk test (SAS Institute, Inc., Cary, NC).

RESULTS AND DISCUSSION

The CRM has been used for about a century to count wildlife populations (Le Cren, 1965). Recently this method has been used in epidemiological studies (LaPorte et al., 1992; McCarty et al., 1993) and is suggested to be the gold standard for counting cases (e.g. prevalence) (Lange et al., 2003a). However, the method can be used to count anything that is countable (Lange et al., 2003). There are a number of criteria that must be applied when using this method. The two-source method requires that each list be independent of each other, although this assumption does not apply if three or more sources are used with a log-linear model (McCarty et al., 1993). Discussion with regulatory officials that compile state lists suggested that there is independence between the state and federal lists, although this must always be tempered with caution when using the two-source method. Observationally, the largest number of sites and missed sites are in counties with the largest population and cost for manufacturing shipments.

Table 1 shows the undercounts (number of sites missed) for state and federal list and the ascertainment (total number of sites that exist according to the CRM calculations). The CI is also provided for the ascertainment. This table is ordered according to the population in each county by sqm. Howard County has a negative number for the state undercount because the total number of state sites was 11 while the ascertainment number was 10.

Spearman rank analysis for undercount numbers and ascertainment

Table 1. Summary data on county/location for estimated number of missed and ascertained hazardous waste sites, population per square mile, cost of manufacturing shipments per year and income per person.

County/ Location	Population	State	Federal	Ascertain-ment/ CI	Manufact Shipment	Income
Baltimore City	8,058	64	106	200(8)	9,822,188	16,978
Montgomery	1,762	5	10	27(3)	3,111,854	35,684
Prince George	1,651	26	43	65(16)	2,008,136	23,360
Baltimore Co.	1,260	71	89	116(40)	6,883,266	26,167
Anne Arundel	1,177	43	61	74(31)	2,703,829	27,578
Howard	983	-1	8	10(5)	1,177,345	32,402
Harford	496	2	12	24(3)	1,274,627	24,232
Calvert	346	0	0	0	404,019	25,410
Carrol	336	0	0	0	716,924	23,829
Frederick	295	5	6	13(2)	1,506,113	25,404
Washington	288	6	14	23(2)	1,924,487	20,062
Cecil	247	2	20	37(4)	678,271	21,384
St. Mary	239	0	0	0	NA	22,662
Wilcomico	224	0	0	0	1,041,565	19,171
Allegheny	176	0	0	0	785,922	16,780
Talbot	126	2	4	7(2)	836,090	28,164
Queen Anne's	109	0	0	0	106,694	23,364
Worcester	98	0	0	0	275,397	22,505
Caroline	93	0	0	0	167,650	17,275
Somerset	76	0	0	0	NA	15,965
Kent	69	0	0	0	189,370	21,573
Dorchester	55	0	0	0	867,153	34,077
Garrett	46	0	0	0	103,452	16,219

Average state population is 542 per sqm; 2000 census
State is undercount for state list
Federal is undercount for CERCLIS list
Manufacture shipments (1997) in $1000
() is confidence interval at 95%
Income is for 1999

suggest that cost of manufacturing shipments and population both have a moderate to good correlation coefficient (Table 2). These correlations are statistically significant. Results suggest that missed HWS are related to the amount of population and manufacturing shipments. More sites are missed in locations where the largest number of manufacturing shipments occur and have larger populations. Areas of low population miss few HWS as compared to heavier populated (urban) areas. When examining population the "cut" point for rural vs. urban areas is suggested to be at the average state population number based on these data. However, use

Table 2. Correlation analysis for number of missed and ascertained sites in each county by state and federal lists and population, manufacturing and income.

Parameters	r^2 value	p value
State with population*	0.584	<0.0001
State with manufacturing+	0.591	<0.0001
State with income	0.104	0.143
Federal with population	0.602	<0.0001
Federal with manufacturing+	0.666	<0.0001
Federal with income	0.071	0.219
Ascertained sites with manufacturing+	0.684	<0.0001
Ascertained sites with population	0.626	<0.0001
Ascertained sites with income	0.295	0.172

*Without Howard County
+Without St. Mary and Somerset Counties

of this value for making this distinction for ascertained and missed HWS requires further evaluation.

Where the manufacturing shipments are less than one billion dollars and the population is less than 400/sqm, very few hazardous waste sites exist, with the exception of Fredrick, Washington and Cecil Counties. Manufacturing cost is suggested to have a stronger correlation than population. Income was not correlated with either federal or state sites and was not statistically significant. This indicates that there is no relationship with income and ascertained or missed HWS. A previous study (Litt and Burke, 2002) of brownfield sites suggested that they are more predominate in poorer areas than those of a higher economic status. However, this investigation was conducted in Baltimore, MD, and city urban areas may have a differing distribution for sites than an entire state where rural and urban areas can be "better" separated.

It is suggested that the best allocation of resources in locating missed HWS is for urban areas. By employing the CRM in determining the number of missed sites on a periodic basis, say every five years, a trend may become apparent as to how well HWS are being identified (located) and where new HWS are occurring (Gioda and Neto, 2002). This will also provide information on how effective a program is on locating unknown HWS and determine the allocation of resources. Analysis of these data overall may also provide information on the public health importance and disease occurrence of unknown HWS in comparison to known sites.

Analysis of this nature can be used to track the effectiveness of HWS activity in a state, regional area or nation. The CRM can be used to evaluate other countable environmental events as well (Lange et al.,

2003). Incorporation of this method into standard practice by environmental and occupational professionals will allow presentation of more accurate numbers. Use of this method in other studies (Lange et al., 2003) would enhance accuracy of counts in those as well (Hillerdal , 1997; Menzie and Wickwire, 2001). This study provides an example as to how the CRM can be effectively used for solving environmental and occupational problems. Findings suggest that more sites are missed in urban areas than rural locations and HWS are best correlated with cost of manufacturing shipment and population. Additional research in these areas is warranted.

REFERENCES

Goida A, Neto FRD. (2002). Exposure to high levels of volatile organic compounds and other pollutants in a printing facility in Rio de Janeiro, Brazil. Indoor Built Environ 11: 302-311.

Hillerdal G. (1997). Pleural plaques: incidence and epidemiology, exposed workers and the general population. Indoor Built Environ 6: 86-95.

Lange JH, LaPorte RE. (2003). Severe acute respiratory syndrome-capture-recapture method should be used to count how many cases of SARS really exist. British Med J 326: 1396.

Lange JH, LaPorte RE, Talbott EO, Chang YF. (2003). Capture-recapture method: the gold standard for incidence and prevalence (Letter). New Zealand Med J 116: U488.

Lange JH, LaPorte RE, Chang YF. (2003a). Exposure to lead and an old way of counting. Environ Health Perspect (correspondence). 111: A510-511.

LaPorte RE, McCarty DJ, Tull ES, Tajima N. (1992). County birds, bees and NCDs. Lancet 339: 494-5.

Le Cren ED. (1965). A note on the history of mark-recapture population estimates. J Animal Ecol 34: 453-4.

Litt JS, Burke TA. (2002). Uncovering the historic environmental hazard of urban brownfields. J Urban Health 79: 464-481.

Maryland Department of the Environment (2002). State Material Hazardous Waste Site List, July, Baltimore, MD.

McCarty DJ, Tull ES, Moy CS, Kwoh CK, LaPorte RE. (1993). Ascertainment corrected rates: applications of the capture-recapture methods. Int J Epidemiol 22: 559-65.

Menzie CA, Wickwire WT. (2001). Defining populations: a key step in identifying spatial and temporal scales. Toxicol Ind Hlth 17: 223-229.

Roney N, Henriques WD, Fay M, Holler JS, Susten SS. (1999). Determining priority of hazardous substances related to hazardous waste sites. Toxicol Ind Hlth 14: 521-532.

Saad SG (2003). Integrated environmental management for hospitals. Indoor Built Environ 12: 93-98.

Sokal RR, FJ Rohlf. (1995). Biometry. WH Freeman, New York. Pg 417 -419.

US Census Bureau. (2003). State and County Quick Facts. http://quickfacts.census.gov/qfd/states/24/24003.html, Washington DC

US Environmental Protection Agency. (2003). Comprehensive Environmental Response Compensation Liability Information System list www.epa.gov/superfund/sites/cursites, Washington, DC