

THEORY AND METHODS

Inter-rater agreement in defining chemical incidents at the National Poisons Information Service, London

I Abubakar, G S Leonardi, N Edwards, N Herriott

J Epidemiol Community Health 2004;**58**:718–722. doi: 10.1136/jech.2002.006650

See end of article for authors' affiliations

Correspondence to: Dr I Abubakar, CDSC Eastern, IPH Cambridge CB2 2SR, UK; ibrahim.abubakar@hpa.org.uk

Accepted for publication 20 October 2003

Background: National surveillance for chemical incidents is being developed in the UK. It is important to improve the quality of information collected, standardise techniques, and train personnel.

Objective: To define the extent to which eight National Poison Information Service specialists in poison information agree on the classification of calls received as "chemical incidents" based on the national definition.

Design: Blinded, inter-rater reliability measured using the κ statistic for multiple raters.

Setting: National Poison Information Service and Chemical Incident Response Service, Guy's and St Thomas's NHS Trust, London.

Participants: Eight specialists in poison information who are trained and experienced in handling poisons information calls and have been involved in extracting information for surveillance.

Results: The overall level of agreement observed was at least 69% greater than expected by chance (κ statistic). Fire and incidents where chemicals were released within a property had a very good level of agreement with κ statistic of 83% and 80% respectively. The lowest level of agreement was observed when no one or only one person was exposed to a chemical (33%) and when the chemical was released into the air (48%).

Conclusion: High levels of agreement were observed. There is a need for more training and improvement in consistency of the data collected by all organisations.

Surveillance for chemical incidents has been an important component of preparedness in the UK, since inception of data collection by the National Focus for Chemical Incidents (NFCI).¹ The importance of surveillance has increased after recent global events and national changes including the publication of the UK's government health protection strategy (Getting Ahead of the Curve).¹ Plans are underway to establish and improve the national surveillance mechanism. The chemical hazards division of the Health Protection Agency is expected to develop a national system of public health surveillance of chemical incidents building on the work of the National Focus for Chemical Incidents. The National Focus for Chemical Incidents in partnership with the Regional Service Provider Units and the Scottish Centre for Infections and Environmental Health (SCIEH) began UK surveillance for chemical incidents with data also being provided by the Ambulance Service Association, National Chemical Emergency Centre, the Maritime and Coastguard Agency, and the Police. This surveillance mechanism will be expected to have standards on timeliness, data quality issues (including case definition and completeness), and outputs. From January to December 2001 a total of 1349 incidents in England, Wales, and Scotland fulfilled the National Focus definition of a Chemical Incident of these 55% were reported by the Chemical Incident Response Service.¹

Public health surveillance can be defined as the ongoing, systematic collection, analysis, interpretation, and dissemination of data regarding a health related event for use in public health action to reduce morbidity and mortality and to improve health.²

This study aims to determine the level of agreement between raters at the National Poisons Information Service, London to improve the quality of data collected for chemical incident surveillance. As there is a national definition, the consistency with which this definition is applied was assessed.

METHOD

A blinded, inter-rater reliability study design with eight observers was used.

"Call sheets" are paper records of inquiries received by the National Poisons Information Service; they are filled out by health professionals, specialists in poison information, or medical staff, at the time of inquiry. In routine practice specialists in poison information select call sheets related to "chemical incidents" and forward to Chemical Incident Response Service for inclusion in the surveillance database. In this study, a sample of 80 call sheets were selected with a wide variety of incidents ranging from those that are clearly not chemical incidents to those that closely fit the national definition of chemical incidents intended for use in surveillance as outlined below. The National Focus for Chemical Incidents defines a chemical incident as *an acute event in which there is, or could be, exposure of the public to chemical substances which cause, or have the potential to cause ill health.*¹

Occupational incidents are included only when members of the public are potentially or actually affected. Hospital staff and emergency services personnel are regarded as members of the public.

The following incidents are included: workplace incidents resulting in off site exposure, for example: a spill in the workplace, resulting in admission to hospital and exposure of hospital staff to the chemicals involved. Other examples include the exposure of people not employed by the company in question, cases of non-intentional carbon monoxide poisoning where more than one person is or might be exposed, unauthorised exposure of the public to CS gas where more than one person is exposed, spills of chemicals in a school laboratory, resulting in the admission of pupils to hospital or exposure of staff or pupils, cases of pesticide over-spray.

Examples of incidents to be excluded from the national surveillance scheme are: occupational exposure covered by

the Health and Safety at Work, etc Act 1974, with no potential exposure to the public, hospital staff, or emergency services, food contamination incidents dealt with under the Food Safety Act 1990, incidents involving radioactive substance, accidental childhood poisoning, drugs and other substances of misuse, suicide attempts involving carbon monoxide, natural gas, and other chemicals.

This definition was presented to the National Poisons Information Service information officers during a training seminar and a copy with examples as outlined above included with the data collection tool for their reference.

The first 80 consecutive calls between 1 April and 30 June 2002 were included. The forms were compiled and data on variables listed in the box were extracted. Eight information officers from the National Poisons Information Service were asked to identify which calls satisfy the definition of a "chemical incident", assuming no one else is extracting incidents for surveillance. They were asked to base their decision on the national definition of chemical incidents and decide which call sheets they would forward to the Chemical Incident Response Service surveillance database. None of the information officers was aware of each other's assessment and call sheets were anonymised.

A database was subsequently established where each record represents a call sheet and the data items in the box were the variables, in addition a variable for each National Poisons Information Service operator participating in the study was created. The National Poisons Information Service operator variable was coded one if that call sheet was classified as a "chemical incident" and zero if not. The proportion classified as incidents and non-incidents by each information officer was summarised. Descriptive characteristics of the 80 call sheets were tabulated and percentages calculated. The κ statistic estimates agreement between raters rating two or more outcomes. The κ statistic is scaled to be zero when the amount of agreement is what would be expected to be observed by chance and one when there is perfect agreement. The κ statistic (two outcomes, multiple raters)³ was calculated to assess the level of agreement between raters. The κ statistics were also calculated within subgroups based on type of incident, location of incident, and chemical involved to enable the identification of the subgroups with a higher level of agreement. The influence of number of people exposed and the presence of health effects on the level of agreement was also examined.

Data items extracted from the National Poisons Information Service call sheets

- Name or identifier of person extracting data.
- Unique identifier for National Focus for Chemical Incidents call sheet.
- Whether call was a "chemical incident" or not.
- Number exposed.
- Presence of health effects.
- Type of incident: airborne release, contaminated land, explosion, deposit, fire, land, malicious act, other, release, spill, unknown.
- Chemical involved:
- Where released to: air, water course/body, water supply, food chain, soil, property.
- Type of site: control of major accident hazards site, residential, industrial, health care, educational, public, agricultural, open space, laboratory, other, not known.

All statistical analysis was carried out using the software Stata (version 6)

RESULTS

Description of calls received

Most calls were about events occurring at home with 42 calls (53.2%) reporting "residential" as site of event, followed by those in which the site was not included in the call sheet with 20 calls (25.3%) and industrial sites (13.9%). There was no difference in the proportion of calls reported as incidents between those with missing data and those with complete information ($p = 0.2$). Fifty three calls sheets reported only one person exposed, 11 calls reported two or more people. Six cases each of spills and deposits were reported and most calls were classified in the "other" category (usually intake of drugs or poisoning). A wide variety of chemicals were reported with over 65 different type of substances reported. Sixty six per cent of calls did not involve any health effect on exposed people while 33.7% (27 calls) reported symptoms or signs after exposure to the chemicals (table 1).

The proportion of calls identified as "chemical incidents" by specialists in poison information ranged from 13.8% to 31%. But most specialists in poison information identified between 22 and 25 calls of 80 as incidents with a median of 22.5 (table 2).

Multiple rater agreement (κ statistic)

The overall level of agreement observed was at least 69% greater than can be expected by chance (κ statistic) (table 3). Fire and incidents where the chemical was released into a property had a very good level of agreement with κ statistic of 83% and 80% respectively. The lowest levels of agreement were observed with incidents involving spills (26%), when no one or only one person was exposed to a chemical (33%), when the chemical was released in to the air (48%), and among incidents with type classified as "other" (mainly drug related calls). Fairly high levels of agreement between raters were also noted in events where no health effect was observed on the exposed (73%) and those occurring at residential sites (73%).

DISCUSSION

The national definition is intended to be pragmatic and sensitive thereby identifying small and large incidents ensuring that data are available for planning, rapid identification of an increase in the number of incidents, and to improve prevention of and preparedness for incidents.

The definition aims to prevent double counting of incidents by excluding those incidents dealt with by other agencies such as the Health and Safety Executive (HSE) but including incidents that chemical provider units would provide advice. This is difficult as incidents, particularly larger ones, will often require a response from a number of different agencies, for example a chemical spill in a hospital may be investigated by the HSE for the purposes of the Health and Safety at Work Act 1974 etc, but the hospital may ask for advice from the chemical provider unit about toxicology or incident management. This type of complexity has resulted in the existing definition and difficulties with interpretation.

One of the key areas to examine in the evaluation of a surveillance scheme is the quality of data collected.² Important aspects of data quality include its completeness, validity (including the validity of the case definition), the quality of training and supervision of persons who complete these surveillance forms, and the care exercised in data management. This study sought to determine the level of agreement between raters and to use various factors to identify issues on which there is lack of agreement between raters. We found a high level of agreement overall but certain

Table 1 Proportion and number of call sheets by type of incident, site, number of people exposed, presence of health effects, and where chemical released to from a total of 80 call sheets

Characteristic	Category	Number (%)
Type of incident	Missing	20 (24.0)
	Air borne release	3 (3.8)
	Explosion	1 (1.3)
	Deposit	6 (7.6)
	Land	2 (2.5)
	Other	32 (40.5)
	Release	1 (1.3)
	Spill	6 (7.6)
	Unknown	9 (11.4)
	Type of site	Unknown
COMAH		1 (1.3)
Residential		42 (53.2)
Industrial		11 (13.9)
Not known		5 (6.3)
Total		79 (100)
Number of people exposed	0	3 (3.8)
	1	53 (66.3)
	2	6 (7.5)
	3 and above	5 (6.3)
	Missing	13 (16.3)
	Total	80
	Health effect	No
Yes		27 (33.7)
Where released to	Not known	30 (37.5)
	Air	8 (10)
	Water body	31 (38.8)
	Water supply	1 (1.3)
	Soil	5 (1.3)
	Property	8 (10)
		1
	Total	80

COMAH, control of major accident hazards, here refers to selected industrial sites that are regulated under European legislation with the same name.

incidents, for instance, where no one or only one person was affected were less straight forward and will be used for further training. The finding of a high level of disagreement within certain types of incidents or incidents with few people affected among a group of highly trained specialists in poison information indicates that a similar level of lack of agreement may exist in other agencies involved in collecting data. These differences are likely to be further compounded by inter organisation differences because of the varied nature of the bodies involved in public health surveillance for chemical incidents. This has implications for further training within National Poisons Information Service and in other centres and organisations.

Another aspect of data quality highlighted by this study is the proportion of records with missing or unknown data on

type of incident, type of site, and where chemical was released. This has implications both for the interpretation of the findings and for future data collection. The completeness of the data collected needs to improve and steps have been taken to achieve this in both the National Poisons Information Service and the Chemical Incident Response Service. The records with missing data may differ from those with complete information. However, this is unlikely to invalidate the results of the study but any conclusions should be interpreted with appropriate caution.

The recording of acute health effects was good, but as in other studies⁴ no record of long term health effects was available. To be able to properly conduct surveillance for chemical incidents, there is a need to measure both short term and long term health effects of exposure. Standardisation of the data collection process nationally will be a reasonable first step. This should include agreement on methods of follow up for all people exposed to chemical agents in this manner. This can be done using record linkage with local hospital data and by updating initial records of surveillance when new developments are noted. Current infectious disease surveillance in the UK benefits from this process where it is possible to notify and denotify a reported case. Consultants in communicable disease control (health protection) can contribute to the chemical incident surveillance process if they are included as an essential part of the national surveillance.

Developments in environmental epidemiology and environmental public health and the creation of chemicals hazards division in the Health Protection Agency is likely to result in improvements in the system. Environmental public health

Table 2 Proportion and number of call sheets identified as a chemical incident by information officers from a total of 80 call sheets

Information officer (identifier)	Number of incidents (%)
One	23 (28.7)
Two	19 (23.7)
Three	25 (31.2)
Four	22 (27.7)
Five	22 (27.7)
Six	11 (13.7)
Seven	24 (30.0)
Eight	23 (28.7)

Table 3 Level of agreement between all raters using κ statistics (overall and by subgroup)

Group		κ Statistic	p Value
Overall		0.6956	0.00
Health Effect	Yes	0.6178	0.00
	No	0.7321	0.00
Number exposed	Less than or equal to one	0.4821	0.00
	Greater than or equals to two	0.5780	0.00
Type of incident	Other	0.3756	0.00
	Fire	0.8322	0.00
	Spill	0.2567	0.0004
Type of site	Missing	0.6348	0.00
	Residential	0.7254	0.00
	Industrial	0.5651	0.00
	Not known	0.6109	0.00
Where released to	Not known	0.7098	0.00
	Air	0.3319	0.00
	Water body	0.6331	0.00
	Property	0.8005	0.00

will benefit from the establishment of a high quality surveillance of at least a similar standard to public health infectious disease surveillance. The role of the public health physician in the management of chemical incidents has been identified in several publications.^{5,6} Environmental public health specialists can contribute by influencing the development of a high quality surveillance system.

One of the strengths of the study was the use of blinding. This was done by ensuring that all raters were not aware of the initial category a call sheet was assigned to or the rating of other raters or any identification information on the call sheet indicating that it has been previously classified by another information officer as a chemical incident.

Ideally, we would have all National Poisons Information Service officers participating in the study, as this could establish a baseline that could be compared with their output after training/discussions or some other intervention aimed at standardising this type of operation. However, because of practical constraints, we were only able to use eight raters. This is however adequate for statistical validity of the analysis.

The κ statistic for agreement between multiple raters was used, this may have a shortcoming because of the assumption that raters should be independent and in real terms as all raters have a similar level and type of training their rating is not truly independent. Agreement analysis with more than

two raters is a complex and controversial subject.⁷ However, even in the case of more than two raters who are not truly independent, κ can still provide an indication of closer agreement for some outcomes relative to others, as in the present case. For values of κ between zero and one, it has been suggested that values between 0.41 and 0.60 indicate moderate agreement, between 0.61 and 0.80 substantial agreement, and above 0.81 almost perfect agreement.⁸

This study has a number of specific advantages and these include:

We have highlighted the need to clarify in operative terms how a definition of chemical incident is implemented, and the difficulties of adopting a definition. The issues are not just having a common definition but having users interpret the definition in an identical fashion—most importantly to avoid systematic differences leading to failure to detect certain incidents or to artefactual increase in the number of incidents detected.

It has been recognised from previous studies that small incidents with no health effects are more likely to be missed compared with major incidents.^{9,10} Although not a specific finding in this study improving the level of agreement between raters will ensure that smaller incidents that fulfil the definition are not missed.

The national surveillance for chemical incidents has several components. This study has permitted a quantitative description of one of the elements of variability in the National Poisons Information Service—Chemical Incident Response Service chemical surveillance system.

We have extracted the National Poisons Information Service call sheets for which the disagreement is largest and used them as material for discussion/training.

Key points

- National surveillance for chemical incidents is currently being developed in the United Kingdom for which a definition has been formulated.
- We found a high level of agreement between specialists in poison information in rating incidents using the national definition especially in incidents involving fires and those in which chemicals were released into a property.
- We also noted low levels of agreement for incidents where no people were affected and those involving releases in to the air.
- The variability in agreement by incident characteristic and type has implications for consistency of the information collected for surveillance.

Policy implications

- Surveillance of chemical incidents needs to take in to account operator behaviour to avoid a situation where inconsistent data are collected. National policy makers should consider the danger of not achieving consistency.
- The need to carry out training nationally to ensure the collection of consistent information that will inform action to prevent chemical incidents.

Concluding remarks

A substantial level of overall agreement between raters with the κ statistic showing agreement beyond what is expected by chance. Various factors were identified leading to either higher or lower level of agreement. This shows the importance of individual rater's judgement when defining and categorising events as chemical incidents, indicating that quality of data produced by National Poisons Information Service is affected by human factors even within a highly trained and skilled pool of operators. An audit process is already in place at National Poisons Information Service, London that focuses on categorisation of calls including chemical incidents and is aimed at improving it. It will be useful to use further training to improve the already satisfactory level of agreement reported here, and extend the investigation/training to other centres as well as regional service provider units to improve the quality of data collected for national surveillance. This is especially important in view of the possibility of different interpretations of the national definition between and within organisations.

ACKNOWLEDGEMENTS

All National Poisons Information Service Specialists in Poison Information who completed the questionnaires, Nicola Batchelor and Jaime Cunningham.

Authors' affiliations

I Abubakar, Communicable Disease Surveillance Centre Eastern, Institute of Public Health, Cambridge and School of Medicine, University of East Anglia, Norwich, UK

G S Leonardi, N Herriott, Chemical Incident Response Service, Guy's and St Thomas's NHS Hospital Trust, London, UK

N Edwards, National Poison Information Service, Guy's and St Thomas's NHS Hospital Trust

Funding: none.

Conflicts of interest: none declared

REFERENCES

- 1 **Health Protection Agency.** <http://www.doh.gov.uk/hpa> (accessed 24 Jun 2002).
- 2 **CDC Atlanta.** <http://www.cdc.gov> (accessed 29 May 2002).
- 3 **Altaye M, Donner A, Klar N.** Inference procedures for assessing interobserver agreement among multiple raters. *Biometrics* 2001;**57**:584–8.
- 4 **Bowen HJ, Palmer SR, Fielder HM, et al.** Community exposures to chemical incidents: development and evaluation of the first environmental public health surveillance system in Europe. *J Epidemiol Community Health* 2000;**54**:870–3.
- 5 **Ayres PJ.** Major chemical incidents—a response, the role of the consultant in communicable disease control and the case of need for a national surveillance-resource centre. *J Public Health Med* 1995;**17**:164–70.
- 6 **Gunnell DJ.** The public health physician's role in chemical incidents. *J Public Health Med* 1993;**15**:352–7.
- 7 **Landis JR, Koch GG.** An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics* 1977;**33**:363–74.
- 8 **Landis JR, Koch GG.** The measurement of observer agreement for categorical data. *Biometrics* 1977;**33**:159–74.
- 9 **Coleman G.** International clearing house for major chemical incidents. Public health and chemical incidents: guidance for national and regional policy makers in the public/environmental health roles. Cardiff: ICHMCI, 1999.
- 10 **Pershagen G.** Environmental public-health surveillance systems for chemical incidents. *Lancet* 2001;**357**:411–12.

THE JECH GALLERY

doi: 10.1136/jech.2004.021337

The National School Fruit Scheme in England



Free school milk was provided in England from 1937 to 1979, when Margaret Thatcher, then Minister for Education, stopped the scheme, earning her the nickname of “Thatcher, Thatcher, the milk snatcher”. The current government’s “5 a day” programme (in England) aims to increase the consumption of fruit and vegetables. As part of this, the National School Fruit Scheme entitles all children aged 4–6 in state schools to a free piece of fruit or vegetable each school day (currently either an apple, banana, pear, or satsuma). The scheme is coming into effect in 2004 and will entail distributing around 440 million pieces of fruit to over 2 million children in some 18 000 schools across England each year.

Mary Shaw

Department of Social Medicine, University of Bristol, Canynge Hall, Whiteladies Road, Bristol BS8 2PR, UK and South West Public Health Observatory; mary.shaw@bristol.ac.uk

Mary Shaw is funded by the South West Public Health Observatory. Thanks to Edward Shaw.