Original article
Scand J Work Environ Health 2000;26(4):317-321
doi:10.5271/sjweh.548

Collection of data for occupational epidemiologic research - results from a survey of European industry
by Rushton L, Betts D

Key terms: data collection; data retention; epidemiologic research; European survey

This article in PubMed: www.ncbi.nlm.nih.gov/pubmed/10994797
Collection of data for occupational epidemiologic research — results from a survey of European industry

by Lesley Rushton, PhD1, Danielle Betts, MSc2

Rushton L, Betts D. Collection of data for occupational epidemiologic research — results from a survey of European industry. Scand J Work Environ Health 2000; 26(4):317–321.

Objectives This study was carried out to provide information on the development of recommendations for the harmonization of record collection and retention by European industry to aid future epidemiologic research.

Methods A postal survey was carried out of 500 companies, stratified by size, in each of the 7 member states of the European Union. Three questionnaires were used respectively focusing on biographic and work history information, health records, and records containing information on exposures. Each questionnaire investigated the data items collected, the data format and location, the length of retention and factors influencing the length of retention.

Results The response after 2 reminders was 39%, 42%, and 33% for the biographic, exposure, and health questionnaires, respectively. Most companies recorded surname, first names, gender, address, postal code, and date of birth, but information on children, place of birth, and nationality were less frequently recorded. Overall, 26% of the companies had an occupational health service, 56% using a contract service, but the percentages varied greatly between countries. Over 85% of the French and Spanish companies carried out routine health surveillance compared with 20% of the companies in Germany and the United Kingdom. Exposure was generally measured for noise (84%), but only 48% of the companies measured chemicals and particulates. Measurements were more common for specific locations (73%) than for particular jobs or individual workers. Approximately 70% of the records were discarded within 10 years, French companies being the most likely to retain all types of records. Legislation was cited as the strongest influence on retention policies.

Conclusions Unless both individuals and industry are made more aware of the importance of record keeping, future occupational epidemiology in Europe will remain limited.

Key terms data collection, data retention, epidemiologic research, European survey.

Work-related illnesses and accidents are costly to European society in both economic and human terms. In Great Britain alone, the total cost of work-related accidents and ill-health in 1990 was estimated to be GBP 6—12 billion, involving 2.2 million people (1). Epidemiologic methods are often used to investigate and identify potential hazards in the workplace, and accurate records of good quality in relation to the work force can play a key role in these studies. They facilitate the enumeration and identification of the relevant workers, enable health outcomes to be identified, and are useful for the estimation of potential exposures.

A 2-year project funded by the European Union (EU) has been carried out to develop recommendations for the harmonization of record collection and retention by industry within the 15 European member states to aid future epidemiologic research. As part of this project a survey was carried out in European industry to ascertain what records are collected, how long records are kept, and what factors influence the retention of records. This paper reports some of the results, focusing on comparisons between the countries surveyed.

Materials and methods

Three questionnaires were developed, 1 focusing on biographic and work-history information held on workers by industry, the 2nd concerning health records, and the 3rd investigating what records were held pertaining to exposures. Each questionnaire included such topics as the

---

1 MRC Institute for Environment and Health, University of Leicester, Leicester, United Kingdom.
2 MRC Institute for Healing Research, University of Nottingham, University Park, Nottingham, United Kingdom.

Reprint requests to: Dr L Rushton, MRC Institute for Environment and Health, University of Leicester, 94 Regent Road, Leicester LE1 7DD, United Kingdom. [E-mail: lr24@le.ac.uk]
data items collected, the data format, the location in which these records are usually kept, the length of time records are retained, and the factors influencing retention. The questionnaires and covering letters were translated into the appropriate language of the member states included in the survey. They were piloted in the United Kingdom (UK) and France.

The postal survey was carried out in 7 member states, namely, France, Germany, Italy, The Netherlands, Spain, Sweden, and the United Kingdom. These countries were chosen to give a spread from north to south across the European Union, to include countries of differing length of EU membership, and to include countries with different proportions of types of industry, for example, heavy manufacturing or agriculture. Three databases were used to select the sample of companies. In the United Kingdom, companies were identified using the Financial Analysis Made Easy (FAME) database (2), which includes all companies with a turnover or shareholder fund exceeding GBP 700 000 or profits exceeding GBP 40 000. The French sample was selected from a database held by the Institut National de Recherche et de Securite (INRS, National Research and Safety Institute of the Prevention of Occupational Diseases and Accidents), which also administered the survey in France. The companies for all the other countries were selected using the REED database, which is a commercial European database.

Samples of about 500 companies were selected for each country, stratified according to company size as small (0—99 employees), medium (100—499 employees) and large (≥500 employees), using the EU definition of small, medium and large enterprises (3). All the questionnaires were sent out with a reply envelope, and 2 reminders, with additional copies of the questionnaires, were sent out to nonrespondents at 3 weekly intervals. In France, Germany, The Netherlands, and Sweden, where access to some form of an occupational health service is universally available for all employees, the biographic and health questionnaires were sent out to all companies. For Italy, Spain, and the United Kingdom a health questionnaire was only sent if a company responded in the biographic questionnaire that it provided some form of occupational health service, and a similar procedure was followed for all countries for the questionnaire related to exposure records.

**Results**

Overall 39% of the biographic questionnaires were returned, as were 42% of the exposure questionnaires, but only 33% of the health questionnaires, although there was some variation between countries. Generally there was a slightly larger proportion of responses from medium-size industries than from small or large industries.

**Biographic records**

Altogether 71% of the biographic records were held in company personnel or human resource departments, 50% being held in salary or finance departments. Italian and UK companies were more likely to hold them in salary departments, and there was a large gradation by size, with small companies much more likely to hold them in salary than personnel departments. Approximately 75% of the companies used both paper and computer records, with the notable exception of France, where only 60% of the companies used both.

Most of the companies always or usually recorded surname (99%), first names (96%), gender (96%), address (99%), postal code (99%), and date of birth (98%), and there was little variation in this respect by country or company size. Marital status was collected by 92% of the companies, but information on children, next-of-kin, maiden name where applicable, place of birth, and nationality were much less frequently recorded.

An identity number and job title were recorded as part of the work history in 85% of the companies, 77% recorded start and finishing dates for each job, 75% recorded location or site, and 64% listed staff grade. Small companies were less likely to record grade. Overall 63% recorded previous employment, the percentage being over 80% for the companies in France and Spain but fewer than 50% for those in Sweden, Italy, and The Netherlands.

**Health records**

Overall 26% of the companies had their own company health service, and 56% used a contracted health service. This figure varied greatly between countries, the percentage for companies having their own occupational health service being 50% for Spain and 54% for the United Kingdom, but only 7%, 10% and 17%,
respectively, for Italy, Germany and Sweden. Preemployment screening was carried out for all employees by 47% of the companies, but 37% did no preemployment screening at all. Spanish (88%), French (86%), and UK (75%) companies were more likely to carry out preemployment screening than such countries as Sweden (45%) and Italy (37%).

Table 1 gives the percentage of companies responding that they "always" or "usually" collected specific health information as part of their occupational health records. The frequency of recording this information was extremely high for the French companies, due to the fact that routine health surveillance is compulsory in France. The frequency of recording by the other countries varied by item of information, although there was a tendency for Spanish companies to be more likely and UK and German companies to be less likely to record it. Altogether 88% of the Spanish companies reported that routine health surveillance was carried out on all employees, whereas only 18% of UK and 20% of German companies gave the same response. Of the clinical investigations inquired about, spirometry (39%) and clinical laboratory measurements (46%) were the most common. Only 21% of the health records were kept in both paper and computer format, and 61% of the companies used paper records only.

**Exposure records**

Of the companies who responded to the exposure questionnaire, 82% reported that they took exposure measurements to comply with legal requirements, 55% used them as part of a company surveillance scheme, and 39% took them as problems arose. The French and Dutch companies were less likely to cite legislation as a reason for taking measurements, and 77% of the Swedish companies had a surveillance scheme.

Exposure measurements were generally taken for noise (84%), but only 48% of the companies reported taking measurements for chemicals and particulates and only about 10% measured ionizing radiation or vibration. The questionnaire included a list of information that might be collected routinely when exposure measurements are made. The companies were also asked whether written documentation existed that described the history of the company processes. Table 2 gives the percentage of companies responding that they always or usually collected this information.

**Length of retention of records**

Table 3 gives the length of time biographic and health records were retained after an employee leaves a company and also the general length of time exposure records were stored. Between 64% and 71% of the records were discarded by the companies within 10 years, but more than 40% of the health and exposure records were kept for less than 5 years. There were large differences between countries, France generally being much more likely to retain all types of records; for example, 68% of the French companies kept biographic records over 10 years or forever. Sweden and The Netherlands were far more likely to discard all types of records within 10 years (81% and 85%, respectively, of biographic records, 89% and 91%, respectively, of health records, and 79% and 83%, respectively, of exposure records).

Companies were asked what influenced both the amount of information they retained and the length of retention. The most important factor appeared to be

| Type of data collected | Companies recording data (%) |
|-----------------------|-----------------------------|
| Location (site and plant) | 89 |
| Date of sample | 87 |
| Agents measured | 81 |
| Units of measurement | 84 |
| Jobs and tasks being sampled | 69 |
| Sampling strategy | 56 |
| Sampling method | 70 |
| Sampling duration | 70 |
| Method of sample analysis | 63 |
| Environmental conditions | 58 |
| Plant process involved | 67 |
| Plant conditions, outputs | 57 |
| Route of exposure | 59 |
| Personal protective equipment | 72 |
| Workers details | 59 |
| Written documentation available | 83 |
| Site plans | 58 |
| Production capacity | 66 |
| Raw materials | 68 |
| Plant products | 65 |
| Process chemicals | 73 |
| Inspection and accident reports | 85 |
| Personal protective equipment | 85 |

| Length of time records are retained. |
|-------------------------------------|
| **Biographic records** | **Health records** | **Exposure records** |
| N | % | N | % | N | % |
|---|---|---|---|---|---|
| 0—5 years | 304 | 22 | 351 | 42 | 186 | 44 |
| 6—10 years | 533 | 42 | 202 | 24 | 116 | 27 |
| ≥11 years | 287 | 20 | 158 | 19 | 82 | 15 |
| Forever and indefinitely | 173 | 12 | 48 | 6 | 30 | 7 |
| Other | 53 | 4 | 71 | 8 | 26 | 6 |
| Missing | 5 | 0 | 12 | 1 | 3 | 1 |
| Total | 1405 | 100 | 842 | 100 | 423 | 100 |

Scand J Work Environ Health 2000, vol 26, no 4 319
Data for occupational research

legislation, particularly for the retention of exposure records, although the French, German, and Italian companies were also likely to cite legislation as having an important influence on the retention of biographic and health records. Company policy on retention was cited by just over 50% of the companies as influencing the retention of biographic and exposure records, but it affected only 27% of the companies with respect to health records.

Discussion

For successful epidemiologic studies in occupational health, information is needed on the population of concern, the health outcomes, and the relevant exposures. Industrial records are, of course, not the only source of this information. For example, health outcome information is often available from external sources, such as death and cancer registers. However, company records can provide a useful starting point, for example, for enumerating a cohort or obtaining details on workplace histories (4).

Industrial records are an example of a secondary data source from which information can be extracted from a record designed and retained for a purpose other than research, for example, to facilitate remuneration. In contrast, face-to-face interviews and clinical examinations are examples of primary data sources. Establishing the completeness and validity of data from secondary sources is thus important. Industrial records may not exist, may be incomplete or inaccessible, or may contain insufficient information from which to, for example, estimate exposures (5). Increasing concern about the quality and existence of industrial records has been reflected by the development of guidelines by some industry-specific bodies and professional associations (6).

This survey has provided useful insight into current record-keeping practices in industry in 7 EU member states. The overall response rates to the 3 questionnaires were low compared with the rates often found in postal surveys of individuals. However, they compare favorably with response rates achieved by other industry-based health-related surveys (7). The fairly even response achieved across the 7 countries and companies of different sizes means that bias is less likely for comparisons between these groups. Nonrespondents in individual-based surveys have been shown to have different characteristics from respondents, particularly, when the admittance of some sort of "bad practice" is required (8). The nonrespondents to this European survey may have poorer record collection and retention practices than those who did respond.

Much of the basic biographic information, such as surname, forenames, date of birth, address and gender, required for an epidemiologic study were routinely collected by the majority of the companies who responded. Other data, such as marital status, information on children, and reproductive outcomes, was less often recorded, so that studies involving these data would require additional research. Knowledge of individual work histories can contribute to the identification of workers exposed to particular hazards. Job title was the most likely to be recorded for workers with dates and location being slightly less likely to be recorded.

A major difference between countries was found in the provision of occupational health services, with companies in such countries as the United Kingdom and Spain being more likely to have their own company service, while countries such as Germany and Sweden used contracted services. However, this finding very much reflects the lack of a general universal occupational health service for all workers in the former and universal provision in the latter. The use of routine health surveillance in countries such as France and Spain again influenced the frequency with which specific items of health information were recorded.

Noise appeared to be the most frequently measured exposure, followed by chemicals and particulates. This trend partly reflects the larger response from heavy manufacturing industry to the exposure questionnaire when compared with the response to the other 2 questionnaires. Measurements were the most likely to be made for specific locations or sites than for individual workers. It is interesting to note from table 2 that data on workers were also less likely to be recorded at the time of taking a sample. Exposure estimates for epidemiologic studies, particularly if being carried out retrospectively, are thus more likely to have to be made for generic situations, for example, for specific job titles, tasks or sites, than for individual persons in a study.

The survey revealed that most records were discarded within 10 years, although countries such as France were more likely to retain them. This early destruction of records, particularly information identifying individual workers, poses considerable problems for historical cohort studies and nested case-referent studies of diseases with long latency periods between exposure and disease diagnosis. Consideration also needs to be given to the loss of records through company liquidation, bankruptcy, or insolvency. Legislation was cited as the most important factor influencing record retention, and the countries citing this influencing factor were likely to have a longer retention time.

Many of the EU health and safety directives indicate the necessity for carrying out health surveillance and keeping records (9). However, the EU countries have implemented these directives in different ways (10), and the results from this survey reflect this difference. A series of structured interviews carried out with UK
researchers explored the "ideal" data needed to carry out occupational epidemiology (11). A workshop held in September 1997 (12) at the end of this 2-year project recommended that a minimum data set on all employees should be collected for epidemiologic research. In particular, data should uniquely identify an individual worker and facilitate linkage to exposure and health effects information. The workshop participants also recommended that there should be a secure chain of custody for these records. The need to make both individual persons and industry more aware of the importance of record keeping for health and safety issues was highlighted. In the absence of more awareness, future occupational epidemiology in Europe may remain limited in its scope.

Acknowledgments

We would like to thank the European Commission for funding this research. Thanks also go to the INRS for administering the French questionnaires, in particular to Dr Pascal Wild and Dr Raymond Vincent, and to Joyce Gilbert, Mary Stevenson and Rachel Skellett at the University of Nottingham for their help with the European survey.

References

1. Davies NV, Teasdale P. The Costs to the British economy of work related accidents and work-related ill health. London: Her Majesty's Stationary Office, 1994.
2. Financial Analysis Made Easy (FAME). London: Bureau van Dijk Electronic Publishing Ltd, 1996.
3. Office for Official Publications of the European Communities. Enterprises in Europe: third report; vol I (Descriptive analysis). Luxembourg: Office for Official Publications of the European Communities, 1994.
4. Pell S. Epidemiological requirements for medical-environment data management. J Occup Med 1978;20:554—6.
5. Bourbonnais R, Meyer F, Theriault G. Validation of self reported work history. Br J Ind Med 1988;45:29—32.
6. Harris RL. Guidelines for collection of industrial hygiene exposure assessment data for epidemiologic use. Washington (DC): Chemical Manufacturers Association, 1993.
7. Harrington JM, Calvert IA. Research priorities in occupational medicine: a survey of United Kingdom personnel managers. Occup Environ Med 1996;53:642—4.
8. Smith C, Nutbeam D. Assessing non response bias: a case study from the 1985 Welsh Heart Survey. Health Educ 1990;5:381—6.
9. Neal AC, Wright FB. The European Communities' health and safety legislation, vol I: the social dimension 1957—1991. London: Chapman and Hall, 1994.
10. Piotet F. Policies on health and safety in thirteen countries of the European Union, vol II: the European situation. Dublin: European Foundation for the Improvement of Living and Working Conditions, 1996.
11. Rushton, Betts DS. Occupational data—the epidemiologists’ dream: interviews with UK researchers. Occup Health Rev 1999;May/June:15—8.
12. Rushton L, Betts DS. Report from a workshop on the development of recommendations for the collection and retention of data by industry for future epidemiological purposes. Luxembourg: European Commission Directorate General V, 1997.

Received for publication: 18 May 1999