Spatial and Spatio-temporal Epidemiology 9 (2014) 37–45

Contents lists available at ScienceDirect

Spatial and Spatio-temporal Epidemiology

journal homepage: www.elsevier.com/locate/sste

Case Study

A comparison between self-reported and GIS-based proxies of residential exposure to environmental pollution in a case–control study on lung cancer

M. Cordioli [a,b,*], A. Ranzi [b], A. Freni Sterrantino [b,c], L. Erspamer [b], G. Razzini [d], U. Ferrari [d], M.G. Gatti [e], K. Bonora [e], F. Artioli [d], C.A. Goldoni [e], P. Lauriola [b]

[a] University of Parma, Department of Bio-Sciences, Parco Area delle Scienze 11/A, 43124 Parma, Italy
[b] Environmental Health Reference Centre, Regional Agency for Environmental Prevention of Emilia-Romagna, Via Begarelli 13, 41121 Modena, Italy
[c] University of Bologna, Department of Statistical Sciences, Via Belle Arti 41, Bologna, Italy
[d] Clinical Trials Office, Cancer Unit of the Carpi General Hospital, Via Guido Molinari 2, Carpi, Modena, Italy
[e] Department of Public Health, Unit of Epidemiology, Local Health Unit of Modena, Strada Martiniana 21, Baggiovara, 41126 Modena, Italy

ARTICLE INFO

Article history:
Received 31 August 2013
Revised 3 March 2014
Accepted 25 April 2014
Available online 9 May 2014

Keywords:
Exposure assessment
GIS
Lung cancer
Questionnaire
Self-reported exposure
Environmental epidemiology

ABSTRACT

In epidemiological studies both questionnaire results and GIS modeling have been used to assess exposure to environmental risk factors. Nevertheless, few studies have used both these techniques to evaluate the degree of agreement between different exposure assessment methodologies.

As part of a case–control study on lung cancer, we present a comparison between self-reported and GIS-derived proxies of residential exposure to environmental pollution.

649 subjects were asked to fill out a questionnaire and give information about residential history and perceived exposure. Using GIS, for each residence we evaluated land use patterns, proximity to major roads and exposure to industrial pollution. We then compared the GIS exposure-index values among groups created on the basis of questionnaire responses.

Our results showed a relatively high agreement between the two methods. Although none of these methods is the "exposure gold standard", understanding similarities, weaknesses and strengths of each method is essential to strengthen epidemiological evidence.

1. Introduction

Lung cancer has been associated with exposure to various environmental risk factors (Alberg et al., 2007; Field and Withers, 2012). Although less significant than smoke and other environmental agents (e.g., occupational exposure to carcinogens, radon, asbestos, etc.) outdoor air pollution is considered a possible risk factor in the etiology of this pathology (Pope III et al., 2011; Raaschou-Nielsen et al., 2011; Turner et al., 2011). Many studies suggest relative risks up to 1.5 for high versus low estimates of exposure to air pollution (Boffetta and Nyberg, 2003).

Various methodologies have been applied in studies on lung cancer to assess human exposure to possible environmental risk factors. These include: (i) self reported exposure (Chan-Yeung et al., 2003; Hosgood III et al., 2010; Hosseini et al., 2009); (ii) comparison between subjects living in urban versus rural areas (Curwen et al., 1954;

* Corresponding author at: Environmental Health Reference Centre, Regional Agency for Environmental Prevention of Emilia-Romagna, Via Begarelli 13, 41121 Modena, Italy. Tel.: +39 059 4226453; fax: +39 059 241312.

E-mail addresses: mik.cordioli@gmail.com (M. Cordioli), aranzi@arpa.emr.it (A. Ranzi), anna.freni@unibo.it (A. Freni Sterrantino), lerspamer@arpa.emr.it (L. Erspamer), g.razzini@ausl.mo.it (G. Razzini), u.ferrari@ausl.mo.it (U. Ferrari), gi.gatti@ausl.mo.it (M.G. Gatti), k.bonora@ausl.mo.it (K. Bonora), f.artioli@ausl.mo.it (F. Artioli), c.goldoni@ausl.mo.it (C.A. Goldoni), plauriola@arpa.emr.it (P. Lauriola).

http://dx.doi.org/10.1016/j.sste.2014.04.004
1877-5845/

Katsouyanni and Pershagen, 1997); (iii) proximity analysis (Garcia-Perez et al., 2009; Simonsen et al., 2010); and (iv) more sophisticated estimates of air-pollution levels (Nyberg et al., 2000; Hystad et al., 2012; Pope III et al., 2011; Raaschou-Nielsen et al., 2011; Turner et al., 2011). Among others, Geographic Information Systems (GIS) are being used with increasing frequency in environmental epidemiology (Nuckols et al., 2004).

One of the main advantages of using GIS versus collecting subjective measures of exposure is the possibility to evaluate personal exposure in large populations. Both self-reported and GIS-derived exposure may be affected by errors, the first mainly due to subjective perception of risks, the latter because of poor data quality. Although none of these methods can be considered the "gold standard" for exposure assessment to atmospheric pollution (Forastiere and Galassi, 2005), comparing different methods is of great interest because it allows the identification of the weakness and strengths of each method and the development of integrated exposure indicators.

Few studies have investigated the relationship between self-reported and modeled exposure to environmental pollution (Cesaroni et al., 2008; Gunier et al., 2006; Heinrich et al., 2005; Migliore et al., 2009). Published studies have focused on exposure to traffic-related pollution with regard to respiratory symptoms like asthma, cough, bronchitis or rhinitis, while none of these studies have referred to lung cancer.

In the present study, we evaluated the association between self-reported exposure to a variety of environmental risk factors and GIS-derived proxies of exposure in a case–control study on lung cancer. In view of the long latency period of this pathology, we devoted particular attention to reconstructing each subject's exposure history. The goals of this analysis were to evaluate the potential of GIS data in exposure assessment compared with self-reported information, and to highlight the importance of using multiple exposure assessment methods in epidemiological studies where "true" measures of personal exposure are not available.

2. Materials and methods

2.1. Study design and population

Data on cancer incidence for the Province of Modena (Northern Italy) in the years 2000–2005 showed a possible cluster for lung cancer in the District of Mirandola, where the Standardized Incidence Ratio (SIR) for males reached the value of 1.26 (CI 95%: 1.13–1.40) (Pirani et al., 2007). A prospective case–control study (the *IDEALE* project) was carried out to investigate the association between environmental risk factors and lung cancer in an area comprising the 9 municipalities belonging to the Mirandola Health District.

A total of 649 subjects were enrolled (case:control ratio of 1:4). Cases were defined as incident events of lung cancer in the period 2009–2010 and controls were coupled on the basis of sex and age. A summary of the main characteristics of the population enrolled is given Table 1.

Table 1
Summary of the main characteristics of the enrolled population.

	n	%
Number of subjects		
Total	649	100
Cases	130	20
Controls	519	80
Sex		
Males	504	78
Females	145	22
Age (years)		
<50	33	5
50–70	262	40
>70	354	55
Smoking habits		
Smoker	110	17
Ex-smoker	351	54
Non-smoker	188	29
Education		
None/Primary school	362	56
Junior high school	148	23
High school	119	18
Degree	20	3

Subjects were interviewed face to face, using a questionnaire designed to collect personal data and information about lifestyle, active and passive smoking habits, food and alcohol consumption, health status, residential and occupational history.

In particular, participants were asked to give information about exposure to environmental pollution at each address of residence since year 1980. The questions, which derived from standardized questionnaire formats (Erspamer et al., 2007; Goldoni et al., 2003; Migliore et al., 2009; SIDRIA, 1997;), were the following:

(1) The zone of residence is predominantly: rural/residential/industrial
(2) The street of residence is: busy/quiet
(3) Do the majority of windows look out directly onto busy roads: yes/no
(4) Are there crossroads or traffic lights within 100 m of the house: yes/no
(5) Do you find dust on windowsills: always or frequently/sometimes/never

2.2. Reconstruction of residential history

Each address reported in the questionnaire was geocoded (coordinate system: UTM32, datum ED50) through record-linkage by street name and street number to the Regional Database (RDB) of the Emilia Romagna Region. Since some of the municipalities in the study area were not included in this database, some addresses were directly geocoded using a global positioning system (GPS); those remaining were geocoded using free web services (Google Maps and Microsoft Bing).

2.3. Exposure assessment

In the whole study area there was only one fixed air-pollution monitoring station, so these data were not usable to differentiate the geographic variability of exposure in