ITALIAN JOURNAL OF PUBLIC HEALTH

Traffic exposure and medication - a GIS based study on prescription of medicines in the Tyrolean Wipptal

Johannes Rüdisser[c], Peter Lercher[1], Armin Heller[2]

[1]Section Social Medicine, Innsbruck Medical University; [2]Department of Geography, University of Innsbruck

Correspondence to: Peter Lercher,Section Social Medicine, Innsbruck Medical University, Sonnenburgstrasse 16, A6020 Innsbruck, Austria. E-mail: Peter.Lercher@i-med.ac.at

Abstract

Background: Environmental health studies often deal with health problems that are influenced by factors with a strong spatial component. However, most analyses of routinely collected health data (cancer register, medication data etc.) are performed at a spatially higher aggregated level. Analyses based on administrative units are frequently subject to confounding by different exposures or other factors and therefore possible effects may remain uncovered or effect estimates distorted.
Methods: A medication prevalence study in the Tyrolean Wipptal, is presented as example for an address based analysis of public health data. To assess whether road or rail exposure is associated with the prescription of medicines the study population was divided in subgroups based on traffic source exposure levels at each address (exposure groups). The medication prevalence proportion for each exposure group was estimated using health insurance data.
Results: Significant differences in medication for various exposure groups could be found. Mainly the railway and main road exposure groups showed higher medication prevalence in some age and medication groups.
Conclusions: Address based spatial analysis of health register data is still an underutilized Public health technique. A persistent problem is the availability of important covariates (education, income etc.) on a more detailed level than municipalities.
This situation might be improved if national statistics agencies provide such statistical data on the basis of the European wide raster system.

Key words: environmental exposure, aggregation bias, environmental health

Introduction

The quality of environmental epidemiology studies and the significance of their results are strongly dependent on the degree of spatial resolution of the utilized data. Propagation of pollutants is determined by physical rules and has a strong spatial component. Environmental epidemiological analyses are often based on data aggregated by administrative units. Aggregation of spatial data frequently results in confounding by different exposures and therefore potential effects are more difficult to detect and in some cases cannot be demonstrated at all. Even the origins of epidemiological research were strongly influenced by the use of precise geo-referenced health data, for example, in 1854 doctor John Snow mapped out all of the cholera cases that occurred around London. By doing this he was able to visualise that the cholera cases were spread around a specific water pump [1]. Snow considered the transmission of the disease through contaminated water to be the most probable cause [2], many years before the pathogen was even discovered and described by Koch in 1883. Snow supported his theory (which he had developed based on more precise studies on earlier outbreaks of Cholera) with what in our time could be called an analogue Geographic Information System (GIS). His analogue-GIS included not only the exact location of all Cholera fatalities but also information about the suppliers of drinking water and all water pumps [3]. Dr. Snow might have never discovered anything new about Cholera if he had only used spatial aggregated data (such as cholera cases per quarter) for his studies.

Modern environmental health studies also deal, in many cases, with health problems that are influenced by factors with a strong spatial component. One can consider the propagation of pollutants, which can often be traced back to a known spatially located source, or the noise induced annoyance along traffic carriers. However, unlike Snow's cholera example, modern environmental health research mostly studies associations that cannot be reduced to a simple cause-and-effect relationship. Health and well-being are influenced by a vast number of different factors. If and how air pollutants or noise affect an individual person, depends on a broad spectrum of different components. Therefore, quantitative estimations of associations between noise or pollutant exposure and health effects are often very difficult. In many cases the number of people studied isn't sufficient in order to statistically prove associations, since the causes for the studied health problems can be diverse and the associations might be weak. In order to study and quantify effects of environmental exposure on human health and well-being the sample size must be very large. This often leads to very expensive and time consuming study designs due to the costliness of data collection.

Public health institutions as valuable data source

Public institutions (hospitals, regional administration, health insurance funds etc.) gather and manage a large quantity of health related data. This data source has huge potential because it is free of selection concerning the population as well as the spatial distribution. Nevertheless, address based analyses of public health data are still rather uncommon. This is due to a variety of reasons: besides the difficulty in accessing such data, due to privacy terms and limiting data politics in different institutions, there were and still are technical problems too. The above mentioned public health data are often not collected and stored in a way that facilitates ongoing systematic and automated processing and the advantage of its completeness can have the disadvantage that one must manage and analyse vast quantities of data which may lead to time consuming and costly data management and analyses procedures [4]. Another problem is that many environmental epidemiology studies which use existing data can only be done on an aggregated level. Health data, such as the occurrence of certain illnesses, are often unavailable on an address basis but only for municipalities, counties, states or other kinds of administrative units and a reliable estimation of disease prevalence are only possible if precise and differentiated information concerning population density and structure is available. In most cases such data are only available in an aggregated form such as population per municipality or district. Therefore many analyses of environmental epidemiology data are limited to relatively huge spatial units such as municipalities, counties, states or similar administrative units [5, 6, 7]. This leads to significant study limitations because sound and air pollution distribution, and hence the exposure of the affected population rarely stops at administrative borders. Various exposure levels are mixed and associations with possible effects are hard (if possible at all) to prove. Furthermore, modern GIS-based methods like geo-statistical analysis or even simple intersections of several spatial layers with unrelated data origins are strongly limited by spatially data aggregation.

Some official providers of statistical population data try to improve the situation by providing data spatially aggregated on the basis of a regular spatial grid. Since the year 2006 the official provider of statistical data in Austria "Statistik Austria" provides standard population data on the basis of a 250m raster [8, 9]. Unfortunately, for many analyses the resolution of these raster is still not detailed enough. Especially in the alpine region where populated areas are often very small and limited to the bottom of valleys, distortions and uncertainties caused by spatial data aggregation can be huge.

A medication study performed in the Tyrolean Wipptal, on the basis of medication data collected by the official health insurance company, will be presented and discussed as an example of an address based analysis of public health and traffic data.

Methodology

The aim of this prevalence study was to assess whether road or rail exposure is associated with the prescription of medicines [10, 11]. The study population was divided into various strata of subgroups (exposure groups) based on individual (address based) estimation of traffic exposure. For these subgroups medication prevalence proportions were calculated and compared. Besides the fact that medication consumption is influenced by various factors such as individual access to the health system or prescription behaviour of physicians, it is still a good indicator for an existing health problem. The study area includes the whole northern Wipptal, which is the whole area between Innsbruck and the alpine Brenner pass. The Wipptal is strongly affected by one of the most important north-south transit routes crossing the alps. The highway, main road and railway run mostly parallel and close to the small and highly populated valley floor. In the Year 2006 the population in the study area was 30,564.

Data about prescriptions came from a specially designed and conducted database query processed by the Tiroler Gebietskrankenkasse (TGKK) - the main public health insurance fund in Tirol - and comprised medication data from the years 2003 to 2005. The information available included: address, age and medication for each year and person, divided in functional groups. Selection of medication groups was done a priori, based on a literature review of health effects that are associated with traffic exposure [12- 16]. The list of studied medication groups is shown in Table 2.

After modifying and adapting the structure of the original data from the TGKK, the data were geo-referenced using the ESRI ArcGIS 9.2 Geo-referencing tool [17] and the official Tyrolean address database.

Exposure Groups

The distance to the nearest visible point of the highway, main road, railway and local road was estimated for all addresses in the study area. Distance to a source of exposure is a commonly used surrogate indicator in environmental medicine [18-20], because individual estimations of specific pollution or noise exposure for a large number of people would in many cases be too expensive or it is only accessible in a reduced and inadequate quality. To reflect the special situation of an alpine valley where terrain structure plays an important role on exposure, especially on noise exposure, the surrogate indicator distance was enhanced by including a visibility analysis. This means that only traffic structures which were visible and not shadowed for instance by a hillside, were considered for distance estimates [11].

All addresses were divided into "exposure groups" based on their distance to the highway, main road, railway and local roads. Addresses with a shorter distance, as defined for each traffic carrier (e.g. <100m), were assigned to the exposure group of the respective carrier. If an address fulfilled the distance definition for more than one carrier it was assigned to the mixed exposure group. Addresses with a distance to the traffic carriers larger than a defined value (e.g. >200 m for railway and highway based on Table 1) were assigned as reference group (Fig. 1). As the cut off values for the exposure groups might influence the results [7, 21] and only few comparable published studies exist, four sampling groups with different cut off levels were applied (see Table 1). This explorative way of data analysis allowed, not only testing of the homogeneity of exposure within each group, but also the sensitivity of the distance value chosen in the four samplings of Table 1. Sampling group 3 turned out to be the most sensitive sampling indicator and results are reported only for this exposure-sampling group.

Figure 1. Classification of addresses to exposure groups based on distance to traffic carriers in the municipality of Matrei am Brenner.

The aggregation of the addresses into exposure groups was necessary because statistical information concerning population is not provided on an address level due to data protection policy. For the predefined exposure-groups population data (age and sex distribution) could be attained from "Statistik Austria" through a special project request [8] based on our own subdivision of addresses.

For all exposure and sampling groups prevalence, odds ratio and 95% confidence intervals were calculated separately for men and women in the age groups: 0-15 years , 15-29 years, 30-49 years, 50-69 years and 70 and older. The odds ratio and 95% confidence intervals were calculated according to Bland & Altmann [22] by a spreadsheet application (MS-Excel 2003) after linkage with the GISdata-base .A 5% level of significance was chosen.

Table 1. Cut off values for exposure and sampling groups.

Exposure type	Sampling 1	Sampling 2	Sampling 3	Sampling 4
Highway (HW)	Distance to HW < 50m	Distance to HW < 100m	Distance to HW < 150m	Distance to HW < 200m
Mainroad (MR)	Distance to MR < 50m	Distance to MR < 50m	Distance to MR < 50m	Distance to MR < 100m
Local roads (LR)	Distance to LR < 50m	Distance to LR < 50m	Distance to LR < 50m	Distance to LR < 50m
Railway (RW)	Distance to RW < 50m	Distance to RW < 100m	Distance to RW < 150m	Distance to RW < 200m
Mixed exposure (ME)	In more then one of the former groups	In more then one of the former groups	In more then one of the former groups	In more then one of the former groups

Results

In 2001, there were 30,322 people living at 8,393 addresses in the study area. Eleven thousand four hundred and ninety-six (37.9%) received one ore more drugs from the studied drug groups that were paid by the TGKK between the years 2003 and 2005. Ten thousand six hundred and eight-one (92.9%) of the insurants could be successfully georeferenced by the provided addresses. Table 2 shows the number of people for each drug group. The most commonly prescribed drugs are antacids followed by four nearly equally large groups (psychosedatives, antiallergic, asthma and hypertension medication).

Table 2. Overall prevalence of prescriptions.

Type of medication	No. in 2003 - 2005	Percentage of population
Antacids	6290	20.76
Antidepressants, Hypnotics and Sedatives	3361	11.09
Antiallergic medication	3244	10.71
Antihypertensives	2970	9.80
Asthma medication	2723	8.99
ENT-drugs (Rhinologica)	1566	5.17
Lipid lowering drugs	1332	4.40
Coronary therapeutics	465	1.54

In comparing the different exposure groups, significant differences in medication odds ratios for the various drug groups were found (Summary in Table 3). For the *railway* exposure group significant elevated medication level for one or more age and sex groups were found for psychosedatives, antihypertensives, antacids as well as antiallergic medications. The age group mostly affected was 70 years and older. For the main road exposure group the medication prevalence proportion for antiallergic medication and antacids were significantly higher.

For the *mixed exposure* group, which is mainly influenced by main road and railway, elevated proportions were also observed for all previously mentioned drug groups. For the *local road* exposure group no increase in medication prevalence proportions was detected.

For the two most common drug groups - antacids and psychosedatives – the complete results for sampling group 3(odds ratio and 95% confidence intervals) are graphically shown in Fig. 2 and Fig 3 (see Table 1).

Discussion

There are many studies about the effect of traffic exposure on the health of affected populations and a considerable number of publications point out that living along busy roads can lead to an elevated health risk [14, 15, 19, 23- 29]. In the literature there is, however, less support for the observed association between railway exposure and elevated medication levels. Therefore, further investigations are required to replicate the results. An analysis based on aggregated data requires cautious interpretation of the results. Furthermore the limited possibility of controlling for covariates means that some uncertainty still remains and therefore prevents a causal interpretation [30]. What stands out is that increased medications associated with traffic exposure (main road and railway) are particularly related to the 70 year and older age group. Reasons for this might be the longer latency period for traffic related health effects to appear or alternatively a higher susceptibility of older people towards both traffic

Table 3. Overall significant associations (OR and 95% CI) with medication type by traffic exposure and age group.

Type of medication	Odds ratio (95% CI)	Traffic exposure	Age group
Antacids	1.56 (1.15-2.12)	mixed exposure	15-29 yrs
	1.44 (1.16-1.80)	mixed exposure	30-49 yrs
	1.42 (1.12-1.80)	main road	30-49 yrs
	1.42 (1.07-1.87)	main road	50-69 yrs
	1.35 (1.13-1.61)	railway	30-49 yrs
	1.95 (1.48-2.57)	railway	70+ yrs
Psychosedatives	2.24 (1.69-2.97)	railway	70+ yrs
Antidepressives	2.12 (1.55-2.89)	railway	70+ yrs
Antihypertensives	2.00 (1.52-2.63)	railway	70+ yrs
Antiallergica	1.43 (1.03-1.98)	mixed exposure	15-29 yrs
	1.33 (1.01-1.74)	mixed exposure	30-49 yrs
	1.33 (1.03-1.72)	railway	15-29 yrs
	1.58 (1.06-2.33)	railway	70+ yrs

Figure 2. Odds ratio and 95% confidence interval for the medication group antacids stratified by age-exposure groups (using sampling 3 according to Table 1). Note: logarithmic scale for the odds ratio and the 95% confidence intervals. Dark bars indicate significant results.

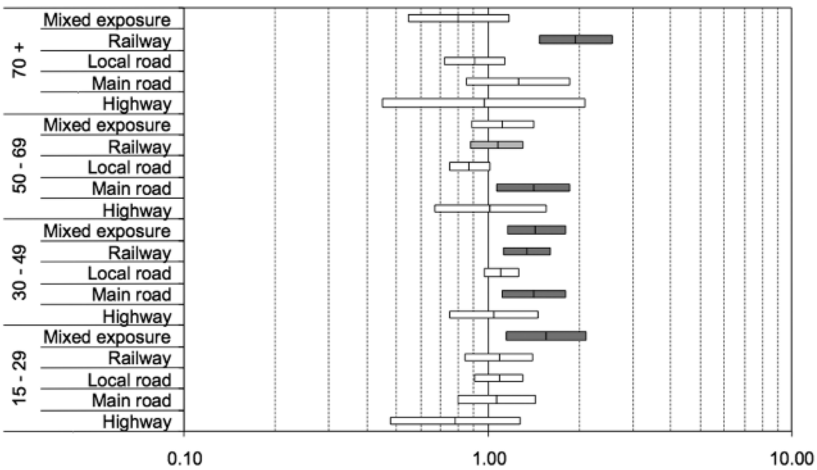
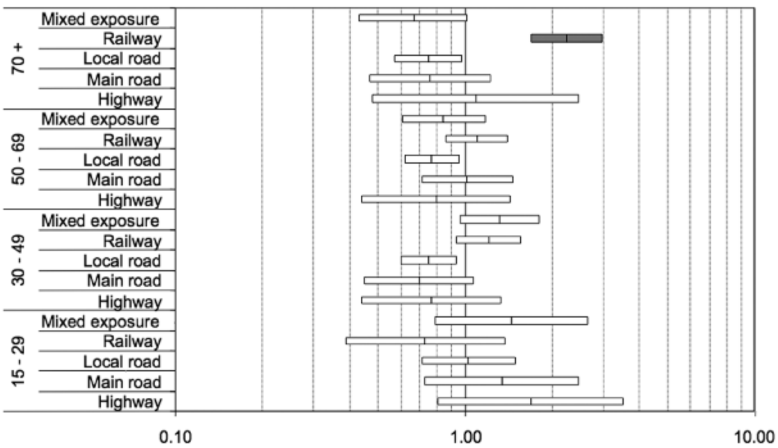


Figure 3. Odds ratio and 95% confidence interval for the medication group psychosedatives stratified by age-exposure groups (using sampling 3 according to Table 1). Note: logarithmic scale for the odds ratio and the 95% confidence intervals. Dark bars indicate significant results.



noise or air pollution. Another hypothesis might be that older people are less mobile and so the exposure estimation on the basis of the living address might correspond more accurately. The fact that no significant increases in medication prevalence proportions were observed for the *highway* exposure group should not lead to the assumption that there may be no effect. There are only a few addresses attributed to being very close (less then 200m) to the highway, so that the number of people in the exposure group highway is very small and confidence intervals are large.

The study results not only show the potential but also the limitations of this kind of address based spatial analysis. The statistical comparison between different exposure groups normalised by age and sex was only possible because of the large number of individuals. Another prerequisite was the availability of statistical data for self defined areas which are supported since the year 2006 by the "Statistik Austria" [8]. The biggest advantage of the analysed data was their completeness (area wide availability of all age groups) and the almost selection free sample (80% of the Tyrolean population is insured by the TGKK). The example presented illustrates the enhanced possibilities of environmental-epidemiology analysis based on exact addresses in contrast to analysis depending on preassigned spatial units such as municipality borders or grids with a low resolution. If analysis of the presented data would have been done on the basis of predefined spatial aggregation, exposure estimates would not have been precise enough to show any associations. An analysis done on the basis of municipalities confirmed this assumption [10]. On the other hand analysis of existing health data on the level of addresses or self defined address based groups comes with a series of problems. Most of the challenges are related to data collection and data availability. The lack of accessibility to address based population data enormously limits the use and the possibilities of analysing existing health register data, as information about population distribution including sex and age related data are the minimum requirement for calculating prevalence estimates. A persistent problem is the lack of important covariates such as education, income, working situation, time spent at home etc., on a more detailed level than municipalities. Possible improvement of this situation might be achieved if national statistics agencies provide statistical data on the basis of a European wide raster system as required by the European Union [9].These data normally have a much better spatial resolution than the common statistical data based on municipalities. Furthermore, they can be combined with exact georeferenced data and so serve as a good database for geo-statistical analysis. Analysis done on the basis of a 250m statistical raster showed promising results.

Finally it must be said that data from public health insurance companies, despite of all its limitations due to missing or hardly accessible reference data, still provides a huge and up till now, not fully utilized potential for further environmental-epidemiology analysis. Routinely collected data of public health care providers offers not only various possibilities for scientific analysis but could also be a cost effective database for the regional environmental health reports required by the WHO [31].

Data privacy protection

As individual prescription data has to be considered as highly sensible data regarding privacy protection a very high standard was set on data protection. Personal data were only used in anonymized forms and results are only published in aggregated forms.

Acknowledgements

We want to thank the following people for cooperation and support to realise often demanding data exchange procedures: Ingrid Kaminger from "Statistik Austria", Mr. Otto Hosp and Mr. Walter Jelinek from Tyrolian GKK, Mr. Manfred Riedl and Manfred Kaiser from the federal state government of Tyrol. Further thanks go to the transnational Brenner Base Tunnel consortium (BBT-SE) for the financial support received.

References

1) Snow J. Report on the cholera outbreak in the parish of St. James, Westminster, during the autumn of 1854. London: Cholera Inquiry Committee, John Churchill,1854.
2) Snow J. On the mode of communication of cholera. London: John Churchill, 1855.
3) Brody H, Rip MR, Vinten-Johansen P, Paneth N, Rachman S. Map-making and myth-making in Broad Street: the London cholera epidemic, 1854. Lancet 2000; 356(9223):64–8.
4) Krieger N. Place, space, and health: GIS and epidemiology. Epidemiology 2003; 14(4):384–5.
5) Nuckols JR, Ward MH, Jarup L. Using geographic information systems for exposure assessment in environmental epidemiology studies. Environ Health Perspect 2004; 112(9):1007–15.
6) Cromley EK. GIS and disease. Annu Rev Public Health 2003; 24:7–24.
7) Tatalovich Z, Wilson JP, Milam M, Jarrett M, McCornell R. Competing definitions of contextual environments. International Journal of Health Geographics 2006; 5:55.
8) Wonka E. Regionalstatistik in Österreich - Von der Tabelle zur räumlichen Analyse und Visualisierung. [Austrian Regional statistic – from the table to spatial analysis]. Salzburg: Selbstverl. des Fachbereiches Geographie und Geologie; 2006. Salzburger geographische Arbeiten; No. 39, 2006.
9) Kaminger I, Meyer W. Neue Raster-orientierte Statistik in Europa. [New raster-oriented statistics in Europe]. In: Strobl J, editor. Angewandte Geoinformatik 2007 [Applied geoinformatic]; Beiträge zum 19. AGIT-Symposium, Salzburg. Heidelberg: Wichmann; 2007:303–8.
10) BBT Brenner Basistunnel-SE. Public Health Study. Innsbruck: Medizinische Universität Innsbruck, Sektion für Sozialmedizin, 2006.
11) Rüdisser J, Lercher P, Heller A. Potenzial und Stolpersteine einer GIS-gestützten Auswertung flächendeckender Medikationsdaten am Beispiel einer Pilotstudie. In: Strobl J, editor. Angewandte Geoinformatik 2007; Beiträge zum 19. AGIT-Symposium, Salzburg. Heidelberg: Wichmann; 2007. p. 637–46.
12) Knipschild P, Oudshoorn N. VII. Medical effects of aircraft noise: drug survey. Int Arch Occup Environ Health 1977; 40(3):197–200.
13) Schulze B, Ullmann R, Mörsted R, Baumbach W, Halles S, Liebmann G et al. Verkehrslärm und kardiovasculäres Risiko: Eine epidemiologische Studie. [Traffic noise and cardiovascular risk: an epidemiological study]. Deutsches Gesundheitswesen 1983; 38:596–600.
14) Lercher P, Kofler W. Behavioral and health responses associated with road traffic noise along alpine through-traffic routes. Sci Total Environ 1996; 189/190:85–9.
15) Lercher P. Environmental noise and health: An integrated research perspective. Environ Intern 1996; 22(1):117–29.
16) Bluhm G, Eriksson A. Aircraft noise exposure and cardiovascular risk among men - First results from a study around Stockholm Arlanda airport. In: Proceedings of the 33rd International Congress and Exhibition on Noise Control Engineering. Prague: The Czech Acoustical Society, 2004.
17) ESRI. Geocoding Rule Base Developer Guide. USA: ESRI, 2003.
18) Verkasalo PK, Kokki E, Pukkala E, Vartiainen T, Kiviranta H, Penttinen A et al. Cancer risk near a polluted river in Finland. Environ Health Perspect 2004; 112(9):1026–31.

19) Hoek G, Brunekreef B, Goldbohm S, Fischer PH, van den Brandt PA. Association between mortality and indicators of traffic-related air pollution in the Netherlands: a cohort study. Lancet 2002; 360(9341):1203–9.
20) Wilkinson P, Elliot P, Grundy C, Shaddick G, Thakrar B, Walls P et al. Case-control study of hospital admission with asthma in children aged 5-14 years: relation with road traffic in morth west London. Thorax 1999; 54:1070–4.
21) Sheppard E, Leitner H, McMaster RB, Tian H. GIS-based measures of environmental equity: exploring their sensitivity and significance. J Expo Anal Environ Epidemiol 1999; 9(1):18–28.
22) Bland JM, Altman GD. Statistics notes. The odds ratio. BMJ 2000; 320(7247):1468.
23) Brunekreef B, Forsberg B. Epidemiological evidence of effects of coarse airborne particles on health. Eur Respir J 2005; 26(2):309–18.
24) Brunekreef B, Holgate ST. Air pollution and health. Lancet 2002; 360(9341):1233–42.
25) Müler-Wenk R. Zurechnung von lärmbedingten Gesundheitsschäden auf den Strassenverkehr. [Evaluation of traffic noise-related health damages] Bern: BUWAL. Schriftenreihe Umwelt; No. 339, 2002.
26) Brauer M, Gehring U, Brunekreef B, de Jongste J, Gerritsen J, Rovers M et al. Traffic-related air pollution and otitis media. Environ Health Perspect 2006; 114(9):1414–8.
27) Hoffmann B, Moebus S, Stang A, Beck E, Dragano N, Möhlenkamp S et al. Residence close to high traffic and prevalence of coronary heart disease. Eur. Heart J. 2006; 27(22):2696–702.
28) Lipfert FW, Wyzga RE, Baty JD, Miller JP. Traffic density as a surrogate measure of environmental exposures in studies of air pollution health effects: Long-term mortality in a cohort of US veterans. Atmospheric Environment 2006; 40:154–69.
29) Morgenstern V, Zutavern A, Cyrys J, Brockow I, Gehring U, Koletzko S et al. Respiratory health and individual estimated exposure to traffic-related air pollutants in a cohort of young children. Occup Environ Med 2007; 64(1):8–16.
30) Greenland S. Ecologic versus individual-level sources of bias in ecologic estimates of contextual health effects. Int J Epidemiol 2001; 30(6):1343–50.
31) Prüss-Üstün A, Corvalán C. Preventing disease through healthy environments: towards an estimate of the environmental burden of disease. Genf: World health Organisation, 2006.