Abstract A panel of 277 children, aged 3–7 years, was used to study the association between air pollution (O₃, SO₂, NO₂, and total suspended particles), meteorological factors (global radiation, maximum daytime temperature, daily averages of vapour pressure and air humidity) and respiratory symptoms. For 759 days the symptoms were recorded in a diary and modelling was based on a modification of the method proposed by Korn and Whittemore (Biometrics 35: 795–798, 1979). This approach (1) comprises an extension using environmental parameters at different time scales, (2) addresses the suitability of using the daily fraction of symptomatic individuals to account for inter-individual interactions and (3) enables the most significant weather effects to be identified. The resulting model consisted of (1) an individual specific intercept that takes account of the population’s heterogeneity, (2) the individual’s health status the day before, (3) a long-term meteorological effect, which may be either the squared temperature or global radiation in interaction with temperature, (4) the short-term effect of sulfur dioxide, and (5) the short-term effect of an 8-h ozone concentration above 60 µg/m³. Using the estimated parameters as input to a simulation study, we checked the quality of the model and demonstrate that the annual cycle of the prevalence of respiratory symptoms is associated to atmospheric covariates. Individuals suffering from allergy have been identified as a group of a particular susceptibility to ozone. The duration of respiratory symptoms appears to be free of scale and follows an exponential distribution function, which confirms that the symptom record of each individual follows a Poisson point-process. This supports the assumption that not only respiratory diseases, but also respiratory symptoms can be considered an independent measure for the health status of a population sample. Since a point process is described by only one parameter (namely the intensity of the point process), it is appropriate for records of respiratory symptoms to identify only one model which covers both the occurrence and duration of symptoms.

Keywords Panel study · Air pollution · Seasonality · Meteorotropy · Ozone

Introduction

Illness prevention is one of the aims of studying the impact of the atmospheric environment on morbidity (Herbarth et al. 1999; Peters et al. 1997; Schwartz 1996; Herbarth 1995; Agocs et al. 1992). There are several methods to record morbidity. For instance, hospital admissions for a specific ICD diagnosis (Spix et al. 1998), reported diagnoses by general practitioners (sentinel system), or the keeping of daily diaries of the health status of a group of persons (panel studies). In the latter the health status is recorded repeatedly by each study participant, which provides a substantial basis for assessing the environmental impact on human health (Schwartz et al. 1994).

One advantage of the panel design is a precisely defined reference population, e.g., a group of pre-school children. Another benefit is that every individual may act as his or her own control and, therefore, results of panel studies can reach higher statistical significance. In practice, panel studies are difficult to conduct, which is probably a reason for their rare application (Peters et al. 1997; Tiittanen et al. 1999). One example is that of Korn and Whittemore (1979), who analysed a panel of 60 asthmatics in Garden Grove, California, for a period of 224 days from 17 November 1974 to 29 June 1975.

In the present paper we describe a panel study of respiratory symptoms in 277 pre-school children during a 2-year period (759 days). Respiratory symptoms are widespread and occur frequently in children, who can be regarded as an especially susceptible population group.
(Goehl 1997). Borsboom and colleagues (1993) observed that adolescents with a positive history of prepubertal respiratory symptoms had consistently lower average ventilation function levels.

Generally the basic analytic objective of panel studies linking symptom status and exposure faces four complications: (1) dependencies among responses of an individual on successive days (auto-correlation), (2) the heterogeneity of the study population (i.e., the individuality of subjects), (3) dependencies among responses of different subjects, and (4) the time-scale over which the exposure influences the response is unknown.

In 1979 Korn and Whittemore published a method that takes account of auto-correlation and heterogeneity. For every individual the probability of showing symptoms is modelled by a logistic approach that is condition- al on the individual’s health status the day before. To allow for heterogeneity, the authors created a two-step approach. In the first step, subject-specific models are identified, i.e., intercept and slope values have to be estimated for every individual. In the second step, all slope values are combined into a common measure for the entire study population.

Dependencies among responses of different individuals have not yet been considered in the literature. However, in the case of respiratory symptoms an inter-individual interaction has to be expected. In the present paper we discuss an attempt to model such mutual respiratory infections in the frame of the Korn and Whittemore approach.

Panel and time-series studies in biometeorology and environmental epidemiology have been used mainly for the detection of short-term effects (e.g., of air pollution), which requires an adjustment for long-term effects like weather and/or season. Temperature and humidity are often taken as surrogates for the prevailing weather conditions (see Schwartz 2000; Delfino et al. 1998). Seasonal, i.e., long-term changes in respiratory complaints are well known and are supposed to be the result of a complex environmental impact (Kim et al. 1996; Donaldson et al. 1999; Seemungal et al. 2000), a phenomenon called meteorotropy.

We systematically search for the appropriate time-scale for each of the environmental covariates and construct a model for the panel data observed in our study. For diagnostic checking we generate panel data from the final model. Comparing both the observed and the simulated data we demonstrate how much of the (seasonal) variation in symptom prevalence is captured by our statistical model. The parameters of the model represent the environmental effects and, finally, we discuss our results in the light of recent literature about air pollution effects.

**Materials and methods**

**Meteorological and air pollution data**

To describe the outdoor air pollution in Leipzig, daily mean concentrations of SO$_2$ (mg/m$^3$), NO$_2$ (mg/m$^3$), total suspended particles (mg/m$^3$), and daily maximums of 8-h moving averages of O$_3$ (mg/m$^3$) were collected at a monitoring station in Central Leipzig (51°20′42′′N, 12°22′52′′E). The meteorological conditions were characterised by daily mean values of relative humidity (%) and vapour pressure (mbar), the daily sum of global radiation (kW/m$^2$), and maximum daily temperature (°C) observed at a meteorological station that is located in Leipzig and is representative for the city area where the studied subjects live. In addition, the length of the day was included in the set of covariates.

**Study population**

This work was part of a larger epidemiological study of the effects of air pollution on the health of children being carried out in Leipzig since 1994. The study comprised 736 subjects and the response to a basic questionnaire was 68.3%. In addition, diaries were kept in a subgroup of 279 of these children who were daily attending one of 17 pre-schools randomly selected from a total of 37. Within the 17 pre-schools, only those children were included in the panel study who had been living in the area for a minimum of 2 years. As the diary was to be kept by the parents, their consent to the study was a precondition for the participation of their child. In our analysis we combined information from the diaries with that of the basic questionnaire and found that the parents of 2 children filled in the diary but failed to complete the basic questionnaire. Thus, the panel study finally comprised 277 children aged 3–7 years (mean 4.5±1.2 years), and their symptoms were recorded for 759 days, from 1 February 1994 until 29 February 1996. This included several children who entered the panel after a delay of some weeks and children who dropped-out prior to the end of the study. The panel attrition is described in Results. In a subset of the panel study group, spirometric measurements were taken and analysed (Fritz and Herbarth 2001).

**Health status and symptom diary**

A symptom diary was developed to be completed by the parents as a continuous daily recording of their child’s well-being or ill-health (Fritz 1993). The items required mainly a yes/no decision and a simple ticking of the appropriate box. For each child, the boxes ticked for (a) sneezing, (b) runny nose, (c) congested nose, (d) cold (white/yellowish nasal discharge), (e) hoarseness, (f) barking cough, (g) dry hacking cough, (h) productive cough and (i) non-productive cough were taken to indicate manifest respiratory illnesses.

Our list of symptoms represents diseases of both the upper and lower respiratory tract, which are indeed often combined in daily practice. Various aspects of the mutual influence of the lower and upper respiratory tract system are described by Hepp et al. (2002) on the basis of anatomical considerations. Analysing epidemiological studies the authors emphasise the concomitant onset of rhinitis, sinusitis, and asthma and give explanations of the pathogenetic mechanisms, the aerogenic transmission and overlapping effects of mediators and inflammatory cells. On the basis of their results it might be appropriate to combine all respiratory symptoms into one outcome variable.

On the other hand, the symptoms in our diary comprise the outcome of several illnesses, ranging from irritations and infections to allergic and asthmatic reactions, which may react differently to environmental factors. Therefore, one general indicator for respiratory symptoms might be misleading and several researchers (such as Corne et al. 2002) partition respiratory symptoms into those of the lower and upper respiratory tract.

It is obviously interesting to compare outcome variables that are differently defined. Therefore we considered the following outcome variables in our study: all respiratory symptoms (a)–(i) were combined into the binary outcome of “general respiratory symptoms” (GRS), the outcome “upper respiratory symptoms” (URS) comprised symptoms (a)–(e); and the symptoms (f)–(i) were combined into the outcome “lower respiratory symptoms”