Eur Respir J 1998; 11: 694–701
DOI: 10.1183/09031936.98.11030694

European Respiratory Journal
ISSN 0903 - 1936

Acute asthma epidemics, weather and pollen in England, 1987–1994

R. Newson*, D. Strachan*, E. Archibald+, J. Emberlin§, P. Hardaker+, C. Collier#

Acute asthma epidemics, weather and pollen in England, 1987–1994. R. Newson, D. Strachan, E. Archibald, J. Emberlin, P. Hardaker, C. Collier. ©ERS Journals Ltd 1998.
ABSTRACT: Recent epidemics of acute asthma have caused speculation that, if their causes were known, early warnings might be feasible. In particular, some epidemics seemed to be associated with thunderstorms. We wondered what risk factors predicting epidemics could be identified.

Daily asthma admissions counts during 1987–1994, for two age groups (0–14 yrs and Š15 yrs), were measured using the Hospital Episodes System (HES). Epidemics were defined as combinations of date, age group and English Regional Health Authority (RHA) with exceptionally high asthma admission counts compared to the predictions of a log-linear autoregression model. They were compared with control days 1 week before and afterwards, regarding seven meteorological variables and 5 day average pollen counts for four species.

Fifty six asthma epidemics were identified. The mean density of sferics (lightning flashes), temperature and rainfall on epidemic days were greater than those on control days. High sferics densities were overrepresented in epidemics. Simultaneously high sferics and grass pollen further increased the probability of an epidemic, but only to 15% (95% confidence interval 2–45%). Two thirds of epidemics were not preceded by thunderstorms.

Thunderstorms and high grass pollen levels precede asthma epidemics more often than expected by chance. However, most epidemics are not associated with thunderstorms or unusual weather conditions, and most thunderstorms, even following high grass pollen levels, do not precede epidemics. An early warning system based on the indicators examined here would, therefore, detect few epidemics and generate an unacceptably high rate of false alarms.
Eur Respir J 1998; 11: 694–701.

*Dept of Public Health Sciences, St George's Hospital Medical School, London, UK. +Observations Division, The Met. Office, Bracknell, UK. §Pollen Research Unit, Worcester College of Higher Education, UK. #Telford Institute of Environmental Systems, University of Salford, UK.

Correspondence: D. Strachan
Dept of Public Health Sciences
St George's Hospital Medical School
Cranmer Terrace
London
SW17 0RE
UK
Fax: 44 1817253584

Keywords: Asthma epidemic
meteorology
pollen
thunderstorm

Received: April 7 1997
Accepted after revision November 11 1997

Although there is a marked seasonal variation in the incidence of asthma attacks, there have been relatively few reports of noticeable "epidemics" of asthma lasting for 1–2 days at a time. However, retrospective exploration of daily rates of health service utilization has identified epidemic days that were not clinically apparent at the time [1]. Repeated outbreaks of severe asthma in Barcelona during the 1980s were attributed to aeroallergen pollution related to unloading of soybeans in the harbour [2], but the cause of other epidemics, such as those in New Orleans [1], remains unclear. Several epidemics of asthma have been reported following thunderstorms in Melbourne [3, 4], Birmingham [5] and London [6, 7]. It has been hypothesized that these may also be explained by exposure to aeroallergens, either disrupted grass pollen [8, 9] or fungal spores [5], which are broadcast by the rapid air movement preceding the storm. Alternative proposed explanations involve the drop in temperature, or increase in humidity, which commonly accompanies a thunderstorm [7, 10].

In view of the suffering, and overloading of emergency services, caused by such epidemics, it would be worthwhile to issue early warnings to hospitals, if enough was known about the causes to predict epidemics in advance. We have shown by analysis of daily time series [11] that a typical severe thunderstorm in an English Regional Health Authority (RHA) is associated with a 25% increase of asthma, and that the excess is greater if preceded by 5 days of high grass pollen. To search for ways of isolating meteorological or other events capable of predicting large epidemics, we have adopted a retrospective "case-control" approach, in which large asthma epidemics were isolated, and compared to "control" days in the same RHA 7 days before and after the epidemics, and also to all nonepidemic days. This approach is complementary to that of our previous paper [11], because it is designed to identify risk factors predicting large epidemics at least some of the time, whereas the previous approach is sensitive to risk factors predicting consistently small excesses, which would not justify early warnings.

Methods

Acute asthma management

Asthma admissions for each of the 14 National Health Service RHAs in England, on each day, were taken from the Hospital Episodes System (HES), which gave counts of asthma admissions for each date, measured from midnight

to midnight, for each District Health Authority (DHA). For each RHA and day, asthma admissions were grouped by patient age into two counts, one for ages ð14 yrs and one for ages Š15 yrs. The time series for the two age groups show a markedly different seasonal pattern, as the younger group has large seasonal peaks and troughs coinciding with the beginnings and ends, respectively, of school terms [12]. The primary data points were, therefore, asthma admissions counts corresponding to combinations of admission date, age group and RHA, grouped into one time series for each combination of age group and RHA, as described previously [11].

Meteorological data

Data were provided by the UK Meteorological Office on sferics (lightning flashes), temperature, rainfall, pressure, humidity and windspeed. For sferic density (sferics per 100 square kilometres per day ($Sf \cdot 100\ km^{-2} \cdot day^{-1}$)), data were collected, for each combination of RHA and day, by methods documented in detail elsewhere [11, 13, 14]. For the other variables, one meteorological station was selected in the territory of each RHA, and data from that station were provided for each day. Days were defined as intervals from midnight to midnight for windspeed data, and as intervals from 09:00 h Greenwich mean time (GMT) to 09:00 h GMT for the other meteorological variables. Temperature was given as the mean between the daily maximum and minimum temperatures in degrees Celsius. Temperature difference for a day was the difference between the mean temperature on that day and on the previous day. Rainfall was given as a daily total in millimetres. Pressure (hecta-Pascals) and humidity (%) were given as values at 09:00 GMT at the beginning of the 09:00–09:00 h day. Windspeed (knots) was given as a daily maximum (1 knot=5.1444 $m \cdot s^{-1}$)

Meteorological data for a day were matched to hospital admissions data for the midnight-to-midnight day on which the meteorological day ended, so that the windspeed data were lagged by 24 h, the sferics, temperature and rainfall were lagged by 15 h, and the pressure and humidity data were lagged by 15–39 h. This was thought to be the most likely match to lead to predictive power, as it allowed time for acute asthma attacks to develop and for patients to travel to hospital. An alternative match was also considered in which meteorological data for a day were matched to hospital admissions data for the day in which the meteorological day began, so that the windspeed data were lagged by 0 h, the sferics, temperature and rainfall data were lagged by -9 h, and the pressure and humidity data were lagged by -9–15 h.

Pollen counts

These data were collated by the Pollen Research Unit, Worcester, UK. Daily mean birch, oak, grass and nettle pollen counts, measured in grains per cubic metre per day ($Gr \cdot m^{-3} \cdot day^{-1}$), were obtained from various stations, one being selected for each of nine of the 14 RHAs in England. Days were defined as beginning and ending at 09:00 h for consistency with the meteorological data. Not all pollen stations possessed a complete pollen record, so gaps were filled in as far as possible by methods documented elsewhere [11]. Pollen counts were entered into the analysis as 5 day moving averages, so that the sferic, temperature and rainfall results for a day were matched with the same hospital admissions as the mean pollen counts for that day and the previous 4 days. This strategy was based on the hypothesis that asthma epidemics are a consequence of the osmotic shock caused by rainfall, which disrupts the reservoir of pollen released over the previous few days, releasing starch granules that are then broadcast by thunderstorm associated wind gusts [8].

Selection of epidemic days

Epidemics were selected as asthma admissions counts (corresponding to combinations of RHA, admission date and age group) which were unexpectedly high, compared to those forecast by a log-linear autoregression model. This model was used by SCHWARTZ *et al.* [15] and originated as a variation on one introduced by ZEGER [16] (see the Appendix, and also our previous paper [11]). Parameter estimation was carried out using a SAS program, available from the authors on request, which uses the procedures GENMOD [17] and NLIN [18] of the SAS/STAT system (SAS System Inc., Cary, NC, USA). A separate model was fitted to the time series of asthma admission counts for each of the two age groups in each RHA in England. These models were nearly the same as those fitted in our previous paper [11], and contained the same number (eight) of autoregressive lag terms, and the same time-dependent x-variates, with the exception of those x-variates describing sferics and pollen counts, which were omitted. The aim of these models was to derive, for each day in each age group in each RHA, a forecast of the asthma admissions count, taking into consideration regional variation, weekly and annual cycles, longer-term trends, and the short-term tendency of an asthma count to be similar to previous asthma counts, but not taking into account the meteorological variables and pollen counts in which we were mainly interested. These forecasts were compared with the actual counts in order to assess the usefulness of the risk factors of interest for predicting counts that are unexpectedly large compared to the forecasts. An epidemic was defined, informally, as an asthma admissions count exceeding the predicted count by at least four coefficients of variation (CVs) (see Appendix for the formal definition).

Selection of control days

Two control days for each epidemic were selected as nonepidemic days in the same series, 1 week before and after the epidemic. This was done to control both for coarse time-dependent background variation and for the "day-of-week effect" familiar in hospital admissions data analysis. The effect of a risk factor (such as temperature) was assessed by calculating the differences between the epidemic-day value and the previous and subsequent control-day values, respectively, and also the signs of these differences (1 for a positive difference, -1 for a negative difference and 0 for a zero difference). The mean of the two differences, and the mean sign for the two differences, were then calculated. These means were in turn averaged

over all epidemics to give a "grand mean difference" and "grand mean sign", with corresponding standard errors and t-tests. Occasionally, an epidemic occurred so soon after the previous one in the same series that the two control sets overlapped in time, and in such a case the two epidemics were not treated as independent, but instead were merged to form a "multi-epidemic episode", entered into the calculations as a single epidemic. The two epidemic-control differences for a multi-epidemic episode were derived, respectively, from the first epidemic and its previous control day and from the last epidemic and its subsequent control day.

Relative risks

Given a categorical risk factor (such as sferic density above or below a threshold), the relative risk of epidemics between two groups was calculated by an odds ratio (OR) method. The relative frequency (odds) of the two groups in the set of epidemics were compared to the corresponding odds in the set of "RHA-days" (combinations of RHA and date, pooled over all RHAs) from which the epidemics in each age group were selected. A confidence interval was defined for the odds of the two groups in the set of epidemics, using an exact binomial confidence interval method [19]. The midpoint and limits of this confidence interval were then divided by the odds of the two groups in the pool of RHA-days, giving a confidence interval for the relative risk.

Results

Table 1 summarizes the 14 RHAs in England, each with the first day of the time series analysed, the number of days with asthma admissions forecasts (from which the epidemics in each age group were selected) and the stations at which meteorological and pollen data were collected. The last day was February 28 1994 for all series, and the first day varied between series, because some RHAs began to report reliable asthma admissions counts earlier than others [11]. For each RHA, the number of days for which there was an autoregressive asthma admissions forecast was the total number of days minus the first eight (for which there are no autoregressive forecasts, as these require eight previous counts). Note that pollen data are available for only nine of the 14 RHAs, and that four RHAs share the London Holloway station as a common pollen data source.

Table 2. – Asthma admission counts+ and epidemics, for the two age groups in the 14 regional health authorities (RHAs) in England

RHA name	Age 0–14 yrs		Age Š15 yrs	
	Admission count	Epidemics n	Admission count	Epidemics n
Northern	7.7 (0.40)	1	6.7 (0.41)	2
Yorkshire	9.4 (0.34)	0	8.4 (0.34)	0
Trent	10.7 (0.34)	2	10.6 (0.35)	4
East Anglian	3.5 (0.57)	1	3.7 (0.55)	5
NW Thames	9.6 (0.37)	3	7.2 (0.40)	2
NE Thames	10.8 (0.36)	2	8.9 (0.41)	7*
SE Thames	9.0 (0.36)	1	7.8 (0.38)	2
SW Thames	6.2 (0.44)	0	4.8 (0.49)	1
Wessex	5.4 (0.46)	1	5.7 (0.46)	1
Oxford	5.4 (0.47)	2	3.9 (0.53)	4
South Western	6.1 (0.47)	1	6.1 (0.46)	3
West Midlands	14.0 (0.30)	3	11.9 (0.32)	2
Mersey	6.9 (0.42)	3	6.6 (0.41)	1
North Western	13.8 (0.30)	2	10.5 (0.33)	0

NW: North West; NE: North East; SE: South East; SW: South West. +: values are mean, and typical coefficient of variation in parenthesis; *: one multi-epidemic episode involved two epidemics on July 22 and 30 1987 in the Š15 yrs age group.

Table 2 summarizes the daily asthma admission counts and gives the number of epidemics, for the two age groups in the 14 RHAs (a list of the epidemic dates is available from the authors on request). The mean counts were calculated directly from all days in the series, including the first eight, which were ineligible for selection as epidemics. The typical coefficient of variation (TCV) is calculated indirectly, by inserting the observed mean into the expression for the coefficient of variation (CV) in Equation (2) of the Appendix. The TCVs varied from 30% in the West Midlands (with the largest mean counts) to 57% in East Anglia (with the smallest mean counts). To qualify as an epidemic under the criteria detailed in the Appendix, an observed count would typically have to be at least 120% in excess of the expected count (more than double)

Table 1. – The 14 regional health authorities (RHAs) in England

RHA name	First day	Days with forecast counts	Weather station	Pollen station
Northern	April 1 1990	1422	Boulmer	-
Yorkshire	April 1 1992	691	Leeming	-
Trent	July 1 1987	2427	Waddington	Derby
East Anglian	July 1 1987	2427	Hemsby	-
North West Thames	April 1 1990	1422	Heathrow	London Holloway
North East Thames	July 1 1987	2427	LWC	London Holloway
South East Thames	April 1 1989	1787	Manston	London Holloway
South West Thames	April 1 1989	1787	Gatwick	London Holloway
Wessex	April 1 1989	1787	Hurn	Isle of Wight
Oxford	April 1 1989	1787	Brize Norton	-
South Western	July 1 1987	2427	St Mawgan	Taunton
West Midlands	April 1 1990	1422	Shawbury	-
Mersey	July 1 1987	2427	Ringway	Chester
North Western	July 1 1987	2427	Aughton	Preston

The last study day was February 28 1994 in all cases. LWC: London Weather Centre.

in the West Midlands, or at least 228% in excess (more than treble) in East Anglia. There was only one multi-epidemic episode (see Methods), involving two consecutive epidemics in North East Thames RHA.

Figure 1 illustrates some points regarding epidemic selection, in the case of the Trent RHA. Figures 1a and 1b show the observed counts for the two age groups as crosses and the forecast counts as lines, which are almost invisible because the crosses are so dense in the neighbourhood of the forecasts (indicating a good fit). Note that the counts for the 0–14 yrs age group are a lot more variable in time than those for the Š15 yrs age group, because of an annual cycle that seems to have maxima at the start of school terms and minima during school holidays [12]. Such predictable annual epidemics had to be eliminated in the search for short-term, possibly weather-related epidemics. Figures 1c and 1d show the standardized residuals, with reference lines at zero (indicating a correct forecast) and at the epidemic threshold. Note that, although the 0–14 yrs age group is more variable, there were only two epidemic days in this RHA once the annual, weekly and long-term effects were removed. The Š15 yrs age group had four epidemics (fig. 1d), two of which correspond to an observed relative excess (ORE) greater than six "CV units" (see Appendix) and counts (fig. 1b) greater than 30. These counts, however, were no larger than those at the highest peaks of annual cycles in the 0–14 yrs age group (fig. 1a).

Figure 2 gives a detailed view of the six episodes in the two age groups in the Trent region, showing the daily count for 1 week before and after each epidemic. Note that if the epidemic takes place on a weekend, there is an apparent tendency for patients to defer their visits to the hospital until the following week. This is another good reason (apart from the weekly cycle) for choosing, as control days, the days 1 week previous and subsequent to the epidemic.

Table 3 gives descriptive statistics for the 11 environmental risk factors under consideration, for all combinations of the 14 RHAs in England and the 2,922 days from January 1 1987 to December 31 1994. The numbers of nonmissing data values are different for different risk factors because of varying availability. For instance, rainfall was available for all days in all RHAs, whereas sferic density was only available in all RHAs from 1990, and pollen counts were only available in nine RHAs, and only for a short time in some of these. Moreover, pollen counts from London Holloway have been "counted four times", because they were matched to all four Thames RHAs (table 1). These descriptive statistics provide a scale against which to compare the observed differences between epidemic and control days.

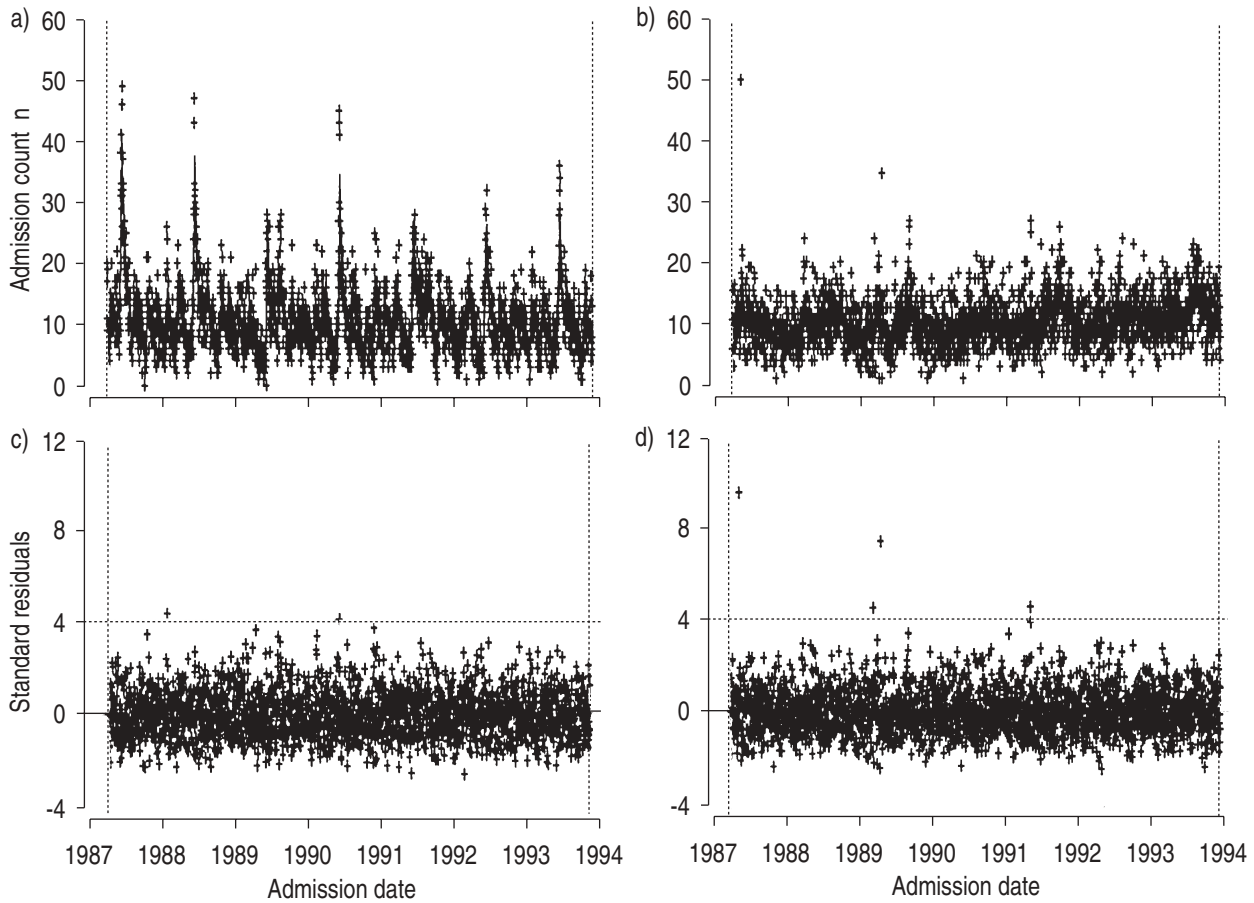

Fig. 1. – Observed and forecast counts and standardized residuals, for the Trent regional health authority, in the two groups. a) Admissions counts in ð14 yrs age group; b) admissions count in Š15 yrs age group; c) standardized residuals in ð14 yrs age group; d) standardized residuals in Š15 yrs age group. Vertical dotted lines are the beginning and end of the observation interval. The years marked on the abscissa refer to April 1 in that year. The horizontal dotted lines (figs. c and d) denote the thresholds for epidemics.

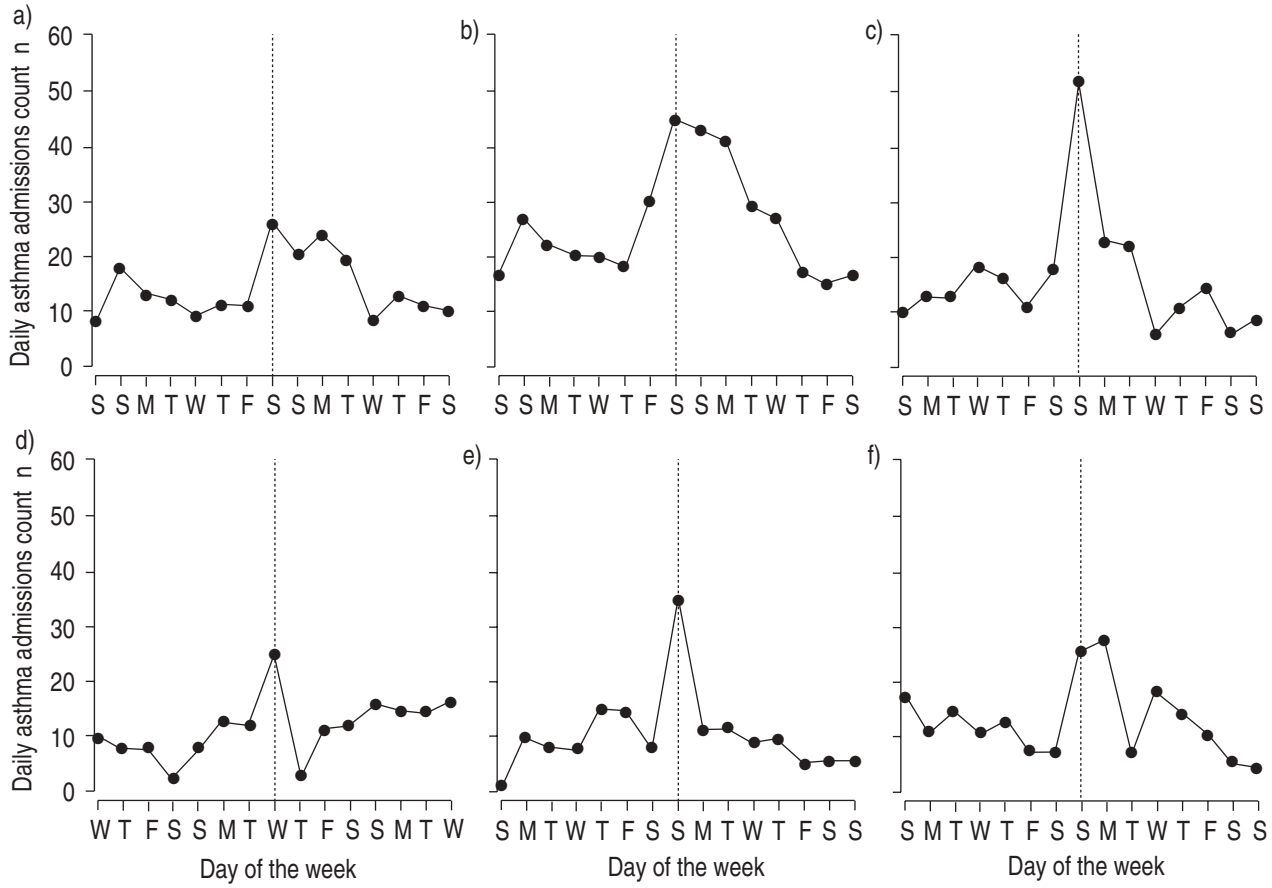

Fig. 2. – Daily counts for 1 week before and after each epidemic in the two age groups in the Trent region. The vertical dotted lines indicate the day of the epidemic. a) episode on April 30 1988 in 0–14 yrs age group; b) episode on September 15 1990 in 0–14 age group; c) episode on August 23 1987 in Š15 yrs age group; d) episode on June 14 1989 in Š15 yrs age group; e) episode on July 23 1989 in Š15 yrs age group; f) episode on August 4 1991 in Š15 yrs age group.

Table 3. – Descriptive statistics for risk factors in all 4,098 combinations of regional health authority and date in England in the years 1987–1994

Risk factor	Values n	Mean	SD	Range
Sferic density Sf·100 km^{-2}·day^{-1}	25578	0.02	0.18	0–8.33
Mean temperature 09:00–09:00 h (°C)	40887	10.19	5.13	-9.00–28.90
Temperature difference °C	40863	0	1.95	-9.50–7.50
Rainfall 09:00–09:00 h mm	40908	1.88	4.09	0–67.50
Relative humidity 09:00 h %	40890	80.07	12.07	25.00–100.00
Pressure 09:00 h hPa	40855	1015.77	11.77	951.40–1047.50
Maximum gust speed 00:00–00:00 h knots	40667	23.12	9.49	0–86.00
Pollen counts Gr·m^{-3}·day^{-1}				
Birch	17143	4.46	23.46	0–351.80
Oak	17019	2.22	11.71	0–188.80
Grass	16953	10.25	29.59	0–364.80
Nettle	16643	9.96	37.70	0–501.60

Sf: sferic; Gr: grains. 1 knot = 5.1444 m·s^{-1}.

Table 4 presents statistics on the epidemic- and control-day values, for the 55 asthma episodes identified, of the 11 risk factors, under the 15 h lag match (see Methods). The mean epidemic-control sign estimates the difference between the respective probabilities of a control value being less than and greater than its corresponding epidemic value, so that, for the example of mean temperature, it is 27% more probable that the control day is cooler than the epidemic day than vice versa. As the risk factors are not normally distributed, any t-tests on the epidemic-control differences must be taken as very approximate, and therefore table 4 contains t-tests on the mean sign, whose distribution is expected to be much more nearly normal (by the central limit theorem). Sferics density is one of the better predictors, although temperature and rainfall also differ between epidemics and control days. Temperatures, for both epidemics and controls, are above the annual averages of table 3, and about 1°C greater (on average) in the epidemics than in the controls. This is consistent with the observation that asthma epidemics often occur during hot spells within warm seasons (as do thunderstorms).

Table 4. – Epidemic-control differences (with mean sign t-tests) for risk factors for asthma episodes

Risk factor	Episodes n	Epidemic mean	Control mean	Difference	Mean sign	Mean sign	
						t-value	p-value
Sferic density Sf·100 km^{-2}·day^{-1}	38	0.39	0	0.39 (0.14)	0.29 (0.07)	3.88	0.0004
Temperature °C	55	13.90	12.68	1.22 (0.45)	0.27 (0.11)	2.52	0.0149
Temperature difference °C	55	-0.44	-0.09	-0.35 (0.31)	-0.05 (0.11)	-0.52	0.6061
Rainfall mm	55	3.12	1.58	1.54 (0.81)	0.20 (0.09)	2.13	0.0381
Humidity %	55	77.68	76.93	0.75 (1.89)	-0.01 (0.11)	-0.08	0.9360
Pressure hPa	55	1016.66	1017.28	-0.61 (1.45)	-0.04 (0.10)	-0.36	0.7186
Gust speed knots	54	21.87	22.31	-0.44 (1.24)	0.03 (0.12)	0.24	0.8133
Pollen counts Gr·m^{-3}·day^{-1}							
Birch	26	1.16	2.73	-1.57 (1.31)	-0.13 (0.08)	-1.77	0.0897
Oak	26	0.31	0.54	-0.23 (0.18)	-0.06 (0.08)	-0.72	0.4777
Grass	26	41.60	28.14	13.45 (8.42)	0.21 (0.11)	1.84	0.0776
Nettle	25	29.73	25.72	4.01 (4.36)	0.18 (0.14)	1.28	0.2142

Values in parentheses are standard errors. For definitions, see legend to table 3.

Table 5. – Relative risk (RR) values of asthma epidemics, for groupings involving high and low positive sferic density and high and low positive grass pollen

Group	RHA-days	Epidemics	RR (95% CI)
Sferics group (all days)			
Zero	18407	26	1.00 (reference)
Low	1669	6	2.55 (0.86–6.32)
High	81	6	52.44 (17.65–130.28)
Pollen groups (all days)			
Zero	7645	6	1.00 (reference)
Low	3730	7	2.39 (0.69–8.61)
High	818	4	6.23 (1.29–26.67)
Pollen group (zero-sferics group)			
Zero	5618	6	1.00 (reference)
Low	2604	5	1.80 (0.43–7.07)
High	555	0	0.00 (0–8.60)
Pollen group (high/low sferics days)			
Zero	383	0	1.00 (reference)
Low	365	2	× (0.20–×)
High	79	4	× (3.20–×)
"Early warning" group (all days)			
No alert	9591	15	1.00 (reference)
Alert*	13	2	98.37 (10.92–422.98)

*: Alert issued if sferics and pollen are both high. High: Š1 Sf·100 km^{-2}·day^{-1} for sferic density or Š50 Gr·m^{-3}·day^{-1} for pollen. RHA: regional health authority; 95% CI: 95% confidence interval. For further definitions, see legend to table 3.

The above analysis was repeated with the -9 h lag (data not shown). The main change observed, when the lag was altered, was that the mean case-control sign for sferic density became nonsignificant in both age groups. Therefore, it seemed that the 15 h lag lost no predictive power when applied to the risk factor of principal interest. The pattern regarding temperature for the -9 h lag was similar to that for the 15 h lag.

From 1990, when automatically collected sferic counts became available, there were 38 epidemics. We have previously defined a high-sferic day in a particular RHA as one with a sferics density at least 1 Sf·100 km^{-2}·day^{-1}, and a low-sferic day as one with a positive sferics density below that threshold [11]. In both age groups, there was a total of six high-sferic epidemics, no high-sferic controls, six low-sferic epidemics, four low-sferic controls, 26 zero-sferic epidemics and 72 zero-sferic controls. If each of the positive sferic groups is entered into a two-by-two table with the zero sferic group, so that positive sferics are tabulated with epidemic-control status, then we find that positive sferics are positively associated with epidemics (Fisher's exact test: p=0.0006 for high sferics, p=0.06 for low sferics).

Sferic-associated epidemics in the two age groups in the 14 RHAs were clustered in time. In fact, even with a threshold as low as 0.07 Sf·100 km^{-2}·day^{-1} (the highest control-day density), they all occurred on five days, August 4 1990, August 26 1991, June 10 1992 and June 10–11 1993. Arguably, therefore, thunderstorm-associated asthma epidemics have only happened once a year since sferic counts became available. Moreover, in each of the 1992 and 1993 epidemics, there was one RHA (North East Thames and South East Thames, respectively) in which there was an epidemic even though the sferic count was zero or very low, respectively. This suggests that thunderstorm-associated asthma epidemics may be more widespread, geographically, than the thunderstorms themselves.

Table 5 reports relative risks of asthma epidemics, for categories and combinations of sferic density and 5 day prior grass pollen measurements, using the thresholds described previously [11]. The third column gives total numbers of RHA-days (see Methods) falling into each category of each grouping. These RHA-days are subsets of those enumerated in the third column of table 1, but the totals

are lower because of missing sferics and pollen values. If RHA-days are grouped by sferic density, moderate thunderstorm days (low sferics) are not significantly over-represented in the epidemics, compared to zero-thunderstorm days, although the data are compatible with a large range of relative risks. Severe thunderstorm days (high sferics) are over-represented compared to zero-thunderstorm days. Similarly, low-pollen days are not significantly over-represented compared to zero-pollen days, but high-pollen days are. When zero-sferic days are grouped by pollen counts, the data are insufficient to measure any effect of grass pollen. However, if nonzero-sferic days are grouped in the same way, we find that thunderstorms following high-pollen spells are at least three times as likely to precede epidemics as thunderstorms following zero-pollen spells, although the difference between low-pollen and zero-pollen spells is not large enough to be measured. Finally, if we imagine that an early warning system has been implemented, issuing alerts whenever a high-sferic day follows a high-pollen spell, then we found that the rare event of an "alert" day is at least 10 times as likely to precede an epidemic as a "no-alert" day, defined as a day on which sferics and prior pollen counts are known but either one or the other is either zero or low. However, of the total of 13 "alert" days observed in the 14 RHAs, representing 26 asthma admission counts in the two age groups, only two were followed by epidemics in one or the other age group. This low number is compatible with an absolute probability of an epidemic in at least one age group of 15% (95% CI, 2–45%), or, in each age group, a probability of an epidemic of 8% (95% CI, 1–25%). These epidemics actually occurred in the 0–14 yrs age group in the North Western RHA on 10 June 1992 and in the ≤15 yrs age group in the North West Thames RHA on 10 June 1993, respectively. Moreover, there are 15 other epidemics for which corresponding pollen and sferics measurements are both available, indicating that, in each age group, 88% of epidemics (95% CI, 64–99%) would not be forecast by the hypothetical early warning system.

Discussion

This report presents a new method for isolating exceptional asthma epidemics out of several series of asthma admission counts. These epidemics can be compared, on the one hand, with control days 1 week before or after in the same age group and RHA, and, on the other hand, with the total pool of RHA-days from which the epidemics in each age group were selected. The first comparison is useful in measuring short-term changes in risk factors which precede an epidemic (such as thunderstorms). The second comparison is useful in measuring the tendency of asthma epidemics to occur preferentially in the presence of longer-term risk factors (such as the pollen season), and in estimating the probability of an epidemic following specified weather conditions. The method is suited for detecting risk factors associated specifically with exceptional epidemics that might cause enough trouble to justify an early warning system. It is complementary to the time-series method defined previously [11], which detected small typical asthma excesses associated with severe thunderstorms and with thunderstorms preceded by a spell of high-pollen days.

The problems associated with identifying epidemics are illustrated in figures 1 and 2. Figure 1 illustrates the presence of background variability. Some of this is attributable to weekly and annual cycles, which lead to regular "epidemics", which are predictable although the causes are not always understood. These "epidemics" must be eliminated in the search for epidemics of the kind of interest here, which are unexpected in the light of present knowledge, and which we would like to find ways to predict. There is also a tendency for there to be runs of consecutive admission counts higher than expected for a given time of year, and log-linear autoregression allows for this, although this might lead to spurious "epidemics" being detected because immediately previous counts happen, by chance, to be lower than expected. The model used here is chosen mainly to detect extremely brief epidemics, such as might be caused by sudden changes in the weather. Epidemics lasting around a week are likely to be attributed to "autocorrelation", although they might have a clinically interesting cause such as a viral epidemic. Figure 2 depicts the appearance of "epidemics" of the kind detected by the present method. Some appear more convincing than others, and it cannot be ruled out that some, at least, might be misidentified.

On average, days preceding epidemics (lagged by 15 h) are warmer by about 1°C than their controls, which are themselves warmer than the average day. Temperature differences, compared to the previous day, are not significantly different in cases and controls. Celenza *et al*. [7], in their analysis of the 2 months surrounding the June 1994 epidemic, claimed that asthma excesses were associated with sudden falls in temperature. This claim is mostly based on their analysis of three-hourly data (detail which we do not have), but a fall in maximum daily temperature was also reported as a highly significant predictor of daily asthma incidence. In a regression analysis of daily data for 2 months surrounding a single spectacular epidemic, any estimates and significance levels are likely to be dominated by contrasts between the epidemic and the general background of surrounding days. Our analysis of epidemics did not detect any spectacular fall in temperature associated with the onset of epidemics, although each epidemic is examined at a much coarser resolution than by Celenza *et al*. [7].

Rainfall, measured at a point source in each RHA, is about twice as high (on average) preceding epidemic days as preceding control days. Pressure, humidity and gust speed, all of which change rapidly at the time of thunderstorms, turn out to be poor immediate predictors. Severe thunderstorms are more likely to precede large asthma epidemics than no thunderstorms. The rare combination of severe thunderstorms preceded by a spell of high pollen counts is an even stronger risk predictor. However, even then, the probability of an epidemic is somewhere between 2% and 45% for either age group, or between 1% and 25% for each age group. Therefore, more than half of the time, this rare combination will be a "false alarm", followed by either a modest excess or no excess at all. The high relative risk, therefore, reflects the extreme rarity of epidemics (around one per thousand admission counts), rather than the high probability of an extreme epidemic following thunderstorms in areas with high prior airborne pollen counts. Conversely, about two thirds (26 out of 38) of epidemics for which sferics data were available (during

1990–1994) were not preceded by thunderstorms (table 5). Possible explanations of these include fungal spores [20], viral epidemics [21] or outdoor pollution, and should be explored in future studies.

The present report, on its own, presents evidence that is significant in a statistical sense, rather than in the sense of justifying a meteorological early warning system based on the risk factors measured here. The pattern that we have observed in this paper, and elsewhere [11], is consistent with the hypothesized mechanism, whereby thunderstorms mobilize aeroallergens from a reservoir of pollen accumulated over previous days [9]. However, the strongest risk factor found here (thunderstorms following high pollen counts) predicts only a small minority of asthma epidemics, with a rate of false alarms too high to be acceptable for the purposes of an early warning system.

Acknowledgements: The authors would like to thank the UK National Asthma Campaign for financial support and the pollen monitoring sites of the European Aeroallergen Network UK for allowing the pollen data to be used in this research.

Appendix: identifying epidemics by log-linear autoregression

The log-linear autoregression model is outlined in the Appendix to our previous paper [11]. The autoregressive residuals (r'_t), defined there, can be divided by their standard deviation to give the standardised residuals (s'_t):

$$s'_t = \frac{r'_t}{\sqrt{\Phi \mu_t}} \quad (1)$$

where Φ is a proportionality constant and μ_t is the predicted count at time t.

These standardized residuals have zero mean and unit variance, and can therefore be used as a standard measure of how incredibly high, or low, a count is, compared to that expected given the observed x-variates and the immediately previous counts. If the individual counts are approximately normally distributed, then the standardized residuals will be approximately standard normal, with zero mean and unit variance.

Epidemics were defined as time points for which the standardized residual exceeded a threshold. This rule defined epidemics in terms of incredibility rather than magnitude. However, if the autoregressive forecast value (μ'_t) is equal to the regressive predicted value (μ_t), as is always the case if there is no autocorrelation, then the observed relative excess (ORE) is given by:

$$\mathrm{ORE} = \frac{y_t - \mu_t}{\mu_t} = s'_t \sqrt{\frac{\Phi}{\mu_t}} = s'_t \times \mathrm{CV}_t \quad (2)$$

where y_t is the observed count at time t, CV_t is the expected Coefficient of variation at time t. The standardized residual, therefore, expresses the ORE approximately in units of the CV, so that a day must have a higher ORE to be accepted as an epidemic where the CV is high (as is typical in a small RHA) than where the CV is low (as is typical in a large RHA). In this paper, a time point was considered to be an epidemic if its standardized residual was at least 4. This is expected with a probability of 0.000032, if the counts are normally distributed, and corresponds, approximately, to an ORE of at least 4 CVs.

References

1. Carroll RE. Environmental epidemiology V. Epidemiology of New Orleans epidemic asthma. *Am J Public Health* 1968; 58: 1677–1683.
2. Anto JM, Sunyer I, Rodriguez-Roisin R, Suarez-Cervera M, Vanquez L. Community outbreaks of asthma associated with inhalation of soybean dust. *N Engl J Med* 1989; 320: 1097–1102.
3. Egan P. Weather or not. *Med J Aust* 1985; 142: 330.
4. Bellomo R, Gigliotti P, Treloar A, Holmes P, Suphioglu C, Singh MB, Knox B. Two consecutive thunderstorm associated epidemics of asthma in the city of Melbourne. *Med J Aust* 1992; 156: 834–837.
5. Packe GE, Ayres JG. Asthma outbreak during a thunderstorm. *Lancet* 1985; 2: 199–204.
6. Venables KM, Allitt U, Collier CG, *et al*. Thunderstorm-related asthma: the epidemic of 24/25 June 1994. *Clin Exp Allergy* 1997; 27: 725–736.
7. Celenza A, Fothergill J, Kupek E, Shaw RJ. Thunderstorm associated asthma: a detailed analysis of environmental factors. *Br Med J* 1996; 312: 604–607.
8. Knox RB. Grass pollen, thunderstorms and asthma. *Clin Exp Allergy* 1993; 23: 354–359.
9. Suphioglu C, Singh MB, Taylor P, *et al*. Mechanism of grass- pollen-induced asthma. *Lancet* 1992; 339: 569–572.
10. Bauman A. Asthma associated with thunderstorms. *Br Med J* 1996; 312: 590–591.
11. Newson R, Strachan D, Archibald E, Emberlin J, Hardaker P, Collier C. The effect of thunderstorms and airborne grass pollen on acute asthma incidence in England, 1990–94. *Thorax* 1997; 51: 680–685.
12. Storr J, Lenney W. School holidays and admissions with asthma. *Arch Dis Child* 1989; 64: 103–107.
13. Lee ACL. An operational system for the remote location of lightning flashes using a VLF arrival time difference technique. *J Atmos Ocean* 1986; 3: 630–642.
14. Lee ACL. Ground truth confirmation and theoretical limits of an experimental VLF arrival time difference lightning flash locating system. *Quart J R Met Soc* 1989; 115: 1147–1166.
15. Schwartz J, Spix C, Touloumi G, *et al*. Methodological issues in studies of air pollution and daily counts of deaths or hospital admissions. *J Epidemiol Community Health* 1996; 50 (Suppl. 1): S3–S11.
16. Zeger SL. A regression model for time series of counts. *Biometrika* 1988; 75: 621–629.
17. SAS Institute Inc., SAS Technical Report P-243, SAS/STAT Software: the GENMOD procedure, Release 6.09. Cary, North Carolina: SAS Institute Inc., 1993.
18. SAS Institute Inc. The NLIN Procedure. *In*: SAS Institute Inc., SAS/STAT User's Guide, Version 6, 4th ed., Volume 2. Cary, North Carolina: SAS Institute Inc., 1989; pp. 1136–1194.
19. Pearson ES, Hartley HO. Biometrika Tables for Statisticians. 3rd ed. (reprinted with additions), Volume 2. Cambridge, Cambridge University Press, 1970; pp. 225–229.
20. Delfino RJ. Daily asthma severity in relation to personal ozone exposure and outdoor fungal spores. *Am J Respir Crit Care Med* 1996; 154; 633–641.
21. Johnston SL, Pattemore PK, Sanderson G, *et al*. The relationship between upper respiratory infections and hospital admissions for asthma: a time-trend analysis. *Am J Respir Crit Care Med* 1996; 154: 654–660.