The temporal pattern and the overall effect of ozone exposure on pediatric respiratory morbidity

Anabela Botelho  
Aida Sá  
José Fraga  
Márcia Quaresma  
Margarida Costa

September 2011

Núcleo de Investigação em Microeconomia Aplicada  
Universidade do Minho
The temporal pattern and the overall effect of ozone exposure on pediatric respiratory morbidity

Anabela Botelho  
(Corresponding Author)  
NIMA and University of Minho  
Campus de Gualtar, 4710-057 Braga, Portugal  
E-mail: botelho.anabela@gmail.com

and

José Fraga¹, Aida Sá¹, Margarida Costa¹, Márcia Quaresma¹  
¹Department of Pediatrics at Centro Hospitalar de Trás-os-Montes e Alto Douro  
Avenida da Noruega – Lordelo, 5000-508 Vila Real, Portugal  
E-mail: zefraga@iol.pt; aida-sa@hotmail.com; anampbc@gmail.com; marciaquaresma@hotmail.com

September 2011

Abstract: Up to now no study has investigated the temporal pattern of children’s respiratory morbidity due to ozone. In the present study, we investigate the temporal pattern and the general effect of ozone exposure on children and adolescents’ respiratory morbidity using data from a particularly well suited area in southern Europe to assess the health effects of ozone. The effects of ozone are estimated using the recently developed distributed lag non-linear models allowing for a relatively long timescale, while controlling for weather effects, a range of other air pollutants, and long and short term patterns. The public health significance of the estimated effects is higher than has been previously reported in the literature, providing evidence contrary to the conjecture that the ozone-morbidity association is mainly due to short-term harvesting. In fact, our data analysis reveals that the effects of ozone at medium and long timescales (harvesting-resistant) are substantially larger than the effects at shorter timescales (harvesting-prone), a finding that is consistent with all children and adolescents being affected by high ground-level ozone, and not just the very frail.

Keywords: non-linear models; distributed lag; delayed effects; public health; respiratory morbidity; children; ozone.

JEL classification: C32; C46; I18; Q53

Acknowledgments: This research was partially funded by Fundação para a Ciência e a Tecnologia (FCT) through the Applied Microeconomics Research Unit (NIMA), award no. PEst-OE/EGE/UI3181/2011.
1. Introduction

Ground-level ozone is a pollutant of growing concern in Europe [1], and children represent the largest subgroup of the population susceptible to the adverse health effects of ground-level ozone concentrations [2], particularly in terms of respiratory diseases [3]. However, there are as yet relatively few epidemiological time-series studies assessing the effects of current European ozone levels on children and adolescents’ respiratory morbidity. Those that exist have been mainly conducted in large urban areas, and the estimated effects tend to be relatively small in magnitude. For example, a recent meta-analysis of European studies provided a summary relative risk of 0.999 for respiratory admissions in children aged 0-14 years per 10 µg/m³ ozone increase [4]. One factor that may account for this finding is exposure misclassification (ie, poor correlation between the commonly used ozone levels measured at fixed sites and personal exposure), an occurrence that tends to cause an underestimation bias in the health effect estimates, particularly in urban areas [1]. In addition, the strong inverse correlation between ozone and fine particles from traffic sources in many cities may confound or conceal the real effects of ozone, or even explain the protective effect of ozone that has been found in some European studies [2], particularly during the winter. Thus, it has been suggested that ambient measurements in warmer places, where people spend more time outdoors, recorded at rural background stations not affected by local traffic pollutants may better represent the actual average population exposure, thereby avoiding the aforementioned confounding factors [5, 6].

Moreover, the existing studies tend to assess the health effects of changes in ozone levels using relatively short timescales. If all extra emergency room visits or hospital admissions in days following higher ozone levels are occurring only a few days early among children already near to acute health events (harvesting, or morbidity displacement), we would expect little association between exposure to higher ozone levels and morbidity counts a few days subsequently [7]. Statistical aspects associated with a morbidity/mortality displacement effect have been discussed by several authors [8], and a conceptual framework for this effect is presented in Figure 1. Area A corresponds to the sum of positive effects on morbidity for highly vulnerable individuals whose illness is being advanced a few days following a high episode of
ozone. Exposure on that earlier day (which produced the short-term morbidity advancement) is then negatively associated with morbidity in subsequent days, when fewer individuals will become ill than otherwise because some of the illnesses have been shifted forward a few days or weeks. This corresponds to the so-called “harvesting” effect, and is captured by area B (sum of negative effects) in the Figure. As pointed out by Zanobetti et al. [8], however, if the increased risk of illness following a high episode of ozone dies out slowly over time, or if there is increased recruitment (by moving people from moderate to severe illness) into the risk pool of vulnerable individuals due to ozone, and this occurs at a slower pace than the effects illustrated in period A, then a period C subsequent to the harvesting period may also be observed where the effects on morbidity are once again positive. The net impact is given by the sum of these areas, and depends on the relative size of each effect.

**Figure 1.** Temporal pattern corresponding to a morbidity displacement effect

Up to now, however, no study has investigated the temporal pattern of children’s respiratory morbidity due to ozone, and only one study has so far analyzed the mortality displacement due to ozone using data for 48 cities in the United States [9]. The authors show that the sum of the ozone effects distributed over three weeks after exposure is larger than the effect estimated using a single day of ozone exposure. Thus, this study indicates that risk assessments using the same or next day after exposure are likely to underestimate, rather than overestimate, the public health impact. Whether the temporal pattern of children’s respiratory morbidity due to ozone causes similar underestimation problems in studies using short timescales is an open question, but the use of longer timescales, presumably resistant to displacement, is
warranted in light of this evidence, along with the findings previously reported in the literature that children tend to express symptoms due to ozone less readily than adults do [10, 11].

In the present study, we investigate the temporal pattern and the general effect of ozone exposure on children and adolescents’ respiratory morbidity expressed in terms of emergency room visits and admissions due to lower airways diseases in a district hospital located in a northeast rural area of Portugal in southern Europe. This is a particularly well suited area to assess the health effects of ozone given that concentrations in southern Europe are higher than in northern Europe and are higher in rural than in urban areas [1, 6]. Ambient ozone measurements in this area may also better represent the population exposure across Europe in the coming decade as ozone concentrations are expected to increase and regional differences in exposure levels to diminish, due to increasing background levels (as currently captured by those measured at rural background stations) and reduced ozone depletion in urban areas [1].

2. Methods

We use daily air quality, meteorological, and hospitalization data for 2004 through 2008 for the rural district of Vila Real in Portugal. The source of the air pollution data is the Portuguese Environmental Agency (www.qualar.org), and includes hourly measurements (in μg/m$^3$) taken at the relevant rural background station of Lamas d'Olo, located in the northeast of Portugal, where ozone (O$_3$) exceedances are very frequent. At this site, pollutants like ozone, sulfur dioxide (SO$_2$), nitrogen dioxide (NO$_2$), particulate matter with aerodynamic diameter less than 2.5 μm (PM$_{2.5}$), and particulate matter with aerodynamic diameter smaller than 10 μm (PM$_{10}$) are monitored since 2004, with hourly acquisition efficiency above 80% [12]. On average, 30% of the total alert threshold exceedances observed by the national monitoring network are registered at this station, reaching more than 40% (and more than 80% of the information threshold exceedances) in 2005 when the highest ozone concentrations in Europe where registered at this station [12, 13]. The annual average ozone concentration at this site is about 100 μg/m$^3$, a value particularly high when compared to the 40-90 μg/m$^3$ average levels detected over the midlatitudes of the Northern Hemisphere [12, 14].
Concerning the monthly distribution of ozone concentrations, the highest values are observed in the summer months (June, July, August and September), but, as already detected by several authors in other different parts of the world, ozone peaks are also observed in the spring, namely in April [12, 15]. As has been previously proposed [2], we restrict our evaluation period to the April-September months. In addition to exclude seasonal influenza activity, these are also the months when children and adolescents spend more time outdoors, and ambient ozone measurements better reflect personal exposure.

Although the levels of other air pollutants tend to be relatively low in the study area, the hourly measurements of SO$_2$, NO$_2$, PM$_{2.5}$, and PM$_{10}$ were also collected and controlled for in the analysis as they have been previously identified as potential effect modifiers of ozone in studies of co-pollutant health effects on children and adolescents [16]. Along with O$_3$, all these hourly measurements were collapsed to daily average values. In order to ensure an equally-spaced and ordered series of the same season for each year, missing data for the air pollutants were imputed following the procedure proposed by Zanobetti et al. [8]. Meteorological measurements were also collected, and controlled for in the analysis, to account for the possible confounding effect of weather on the relationship between ozone and morbidity. The weather measurements are provided by the National Meteorological Service for the study area (Vila Real district), and include daily values for average temperature (in °C), and relative humidity.

The study group consists of all children and adolescents under the age of 18 who resorted to the pediatric emergency room service of the hospital center known as Centro Hospitalar de Trás-os-Montes e Alto Douro (CHTMAD), which serves as the health reference center for the entire district, during the period between 1 April – 30 September for the years 2004 – 2008, and were diagnosed with lower respiratory tract diseases (ie, any infectious or inflammatory disease of the lower airways, including pneumonia, bronchiolitis and asthma exacerbation; all upper airway respiratory disease and trauma pathology were excluded). The daily counts of hospital emergency room visits and daily counts of hospital admissions under the referred conditions were obtained through consultation and analysis of the emergency charts and clinical processes, on paper or digital format, using the Medical IT Support System (SAM) of the hospital center during the selected period.
2.1. Statistical Methods

Estimation of the effects of environmental factors on health outcomes must take into account the (i) nature of the dependent (response) variable; (ii) the potential nonlinear relationship between the dependent variable and the covariates (predictors); and, (iii) the potential delayed effects of the covariates on the dependent variable. In this study, the outcomes of interest are daily counts of hospital emergency room visits and daily counts of hospital admissions, both of which can only take values limited to the nonnegative integers. This suggests that the data is generated by a Poisson process. However, estimation of a Poisson regression model is only appropriate if the data is equidispersed (the variance is equal to the mean function), a feature that is commonly violated in count data. When the data is overdispersed (variance larger than the mean), as in the current application, statistical inference from the Poisson regression is incorrect as it fails to account for the excess variation around the model’s fitted values [17]. This issue can be addressed through the estimation of a so-called quasi-Poisson model which basically consists in estimating an additional “dispersion parameter” from the data, and using it to adjust inference for overdispersion [17].

In their standard formulation, these models assume a linear relationship between the dependent variable and the predictors (or some other pre-defined mathematical function, such as a polynomial, etc.). In environmental epidemiologic studies, however, the relationship between the response variable and some of the predictors is expected to be nonlinear in ways that cannot be easily defined a priori. The generalized additive models (GAM) proposed by Hastie and Tibshirani [18] specify the predictor as an additive function of a transformation of the original predictors, and the related parameters can be estimated within the generalized linear models’ framework using a quasi-Poisson family of distributions. These models are quite flexible, allowing some predictors to be included in the model as smooth functions and some others as linear functions as determined by the degrees of freedom chosen for the respective predictors, and have, therefore, been systematically used in studies investigating the effects of temperature and air pollution on health.

In addition to possible nonlinearities in the dimension of the predictors, environmental epidemiologic studies must take into account the possible time lags between changes in the predictors and their effects on health. A simple approach to take the temporal dimension of the predictors into account consists in introducing as
predictors the moving averages of the environmental exposures up to a certain lag. Such an approach has been extensively used in environmental epidemiologic studies, and is a special case of the so-called distributed lag models (DLM). In their basic formulation, DLMs specify the response variable at any time $t$ as an additive function of the predictors measured not only at time $t$ but also at different lags. An estimate of the overall effect of a unit change in any given predictor is then given by the sum of the estimated coefficients over the whole lag period considered. However, if the predictor does not change much or moves fairly regularly, this specification may suffer from severe collinearity, resulting in large estimated standard errors. One popularly used method to alleviate the collinearity problem consists in specifying an Almon lag scheme [19]. Rather than attempting to estimate directly each of the current and lagged coefficients, this approach assumes that each coefficient can be approximated by a suitable degree polynomial in the lag number so that the curve along lags follows a certain smooth function.

These two types of models, the GAM and the DLM, can be combined giving rise to generalized additive distributed lag models (GADLM) as proposed by Zanobetti et al. [8]. In their formulation, the response variable (assumed to follow a Poisson distribution) is modeled as an additive function of smoothed predictors plus their polynomial distributed lagged effects. Although incorporating both nonlinearities in the dimension of the predictors and their temporal dimension, this approach keeps these effects separated. Recently, Gasparrini et al. [20] developed the so-called distributed lag non-linear models (DLNM) that allow nonlinear dependencies and lagged effects to be modeled simultaneously in quite flexible ways. The mathematical formulation of DLNM is quite complex [20, 21], but it relies on a simple concept: a known set of transformations (basis functions) of the original predictor are independently set to describe the shape of the relationship in each dimension, and these functions are then combined to generate a cross-basis function which is a bi-dimensional space of functions describing simultaneously the shape of the relationship along the space of the predictor and along its temporal dimension.

In this study, we apply the DLNM framework recently implemented within the statistical software R [21] to assess the effects of daily ozone measurements on daily counts of hospital emergency room visits and daily counts of hospital admissions, while controlling for weather effects, a range of other air pollutants, and long and short term patterns.
3. Results and Discussion

Table 1 presents descriptive statistics for the variables used in the analysis. During the period between 1 April – 30 September for the years 2004 – 2008, there were 1952 hospital emergency room visits, and 350 hospital admissions, due to lower respiratory tract diseases.

The maximum number of daily hospital emergency room visits (14), and the maximum number of daily hospital admissions (6) were both registered in year 2005. This was also the year that registered the highest number of days in which average daily ozone concentrations exceeded the current EU target value for ozone concentrations (120 µg/m$^3$; Directive 2008/50/EC). In fact, this target value was exceeded 88 times in 2005, corresponding to 48% of the 183 observed values. Overall, the target value was exceeded in 23% of the 915 days under observation.

Table 1. Descriptive Statistics

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Lower respiratory diseases</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ER visits per day</td>
<td>2.13</td>
<td>2.00</td>
<td>0.00</td>
<td>14.00</td>
<td>915</td>
</tr>
<tr>
<td>Hospital admissions per day</td>
<td>0.38</td>
<td>0.71</td>
<td>0.00</td>
<td>6.00</td>
<td>915</td>
</tr>
<tr>
<td><strong>Pollutants</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$O_3$(µg/m$^3$)</td>
<td>104.30</td>
<td>27.10</td>
<td>39.38</td>
<td>235.73</td>
<td>915</td>
</tr>
<tr>
<td>$PM_{2.5}$(µg/m$^3$)</td>
<td>10.66</td>
<td>9.35</td>
<td>0.50</td>
<td>97.21</td>
<td>915</td>
</tr>
<tr>
<td>$PM_{10}$(µg/m$^3$)</td>
<td>21.74</td>
<td>14.81</td>
<td>0.40</td>
<td>184.96</td>
<td>915</td>
</tr>
<tr>
<td>$SO_2$(µg/m$^3$)</td>
<td>2.61</td>
<td>3.00</td>
<td>0.00</td>
<td>20.04</td>
<td>915</td>
</tr>
<tr>
<td>$NO_2$(µg/m$^3$)</td>
<td>2.74</td>
<td>2.08</td>
<td>0.00</td>
<td>22.88</td>
<td>915</td>
</tr>
<tr>
<td><strong>Weather</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temperature ($^o$C)</td>
<td>18.35</td>
<td>4.88</td>
<td>5.00</td>
<td>30.80</td>
<td>915</td>
</tr>
<tr>
<td>Relative humidity (%)</td>
<td>63.85</td>
<td>15.41</td>
<td>23.00</td>
<td>96.00</td>
<td>915</td>
</tr>
</tbody>
</table>

Pearson correlation coefficients between the considered air pollutants and meteorological variables are presented in Table 2. The results show that $O_3$ is positively correlated with $PM_{2.5}$, $PM_{10}$, and $NO_2$ at better than the 0.05 significance level; the correlation of $O_3$ and $SO_2$ is small in magnitude and statistically insignificant. All the pollutants are substantially correlated with the weather variables at better than the 0.05 significance level: they are positively correlated with temperature and negatively correlated with humidity.
Table 2. Pearson correlation coefficients

<table>
<thead>
<tr>
<th></th>
<th>O₃</th>
<th>PM₂.₅</th>
<th>PM₁₀</th>
<th>SO₂</th>
<th>NO₂</th>
<th>Temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>O₃</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PM₂.₅</td>
<td>0.4688*</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PM₁₀</td>
<td>0.4457*</td>
<td>0.7888*</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SO₂</td>
<td>0.0023</td>
<td>0.0288</td>
<td>0.1489*</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NO₂</td>
<td>0.2460*</td>
<td>0.4878*</td>
<td>0.5088*</td>
<td>0.3636*</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>Temperature</td>
<td>0.3575*</td>
<td>0.4056*</td>
<td>0.4315*</td>
<td>0.1834*</td>
<td>0.3300*</td>
<td>1.000</td>
</tr>
<tr>
<td>Rel. humidity</td>
<td>-0.4458*</td>
<td>-0.2915*</td>
<td>-0.3106*</td>
<td>-0.1676*</td>
<td>-0.1774*</td>
<td>-0.5282*</td>
</tr>
</tbody>
</table>

As explained above, the effect of ozone on daily counts of hospital emergency room (ER) visits and daily counts of hospital admissions (HA) was estimated within the DLNM framework using a generalized linear model with quasi-Poisson family to account for overdispersion (the estimated dispersion parameter is 1.33 and 1.16 in the ER visits model and in the HA model, respectively). The relationship with ozone was modeled through a natural cubic spline with 4 degrees of freedom and boundary knots located at the range of the observed values. The relationship in the space of the other air pollutants and weather variables was modeled through a natural cubic spline of 4th order with boundary knots located at the range of their respective observed values. In all cases, the lagged effect was specified up to 30 days of lag with a 5th degree polynomial function. In addition, indicator variables for year, month of the year, season of the month (spring and summer), an interaction term between month and season, and day of the week, were included in the model to control for short-term confounders and seasonality. The specifications for the degrees of freedom in each dimension were chosen so as to minimize the quasi-Akaike Information Criterion [20]. A maximum lag of 30 days was allowed in order to examine the temporal pattern of children’s respiratory morbidity due to ozone.
Figure 2. Plot of RR at specific lags (Reference at 120 µg/m³)

Figure 2 shows the lag-specific effects for a 10-unit increase in O₃ compared with a reference value of 120 µg/m³, the current EU target value for ozone concentrations. The top panel shows the relative risk (RR) concerning emergency room visits along with 95% confidence interval, and the bottom panel shows the RR concerning...
hospital admissions. Importantly, in each case the figure shows the basic features of the temporal pattern postulated in Figure 1.

A short-term advancement is observed for emergency room visits at lags 0-3, and at lags 0-2 for hospital admissions. A harvesting period is also observed at lags 4-10 and at lags 3-9 for emergency room visits and hospital admissions, respectively. In each case, the harvesting period is followed by a long period where positive effects tend to persist, suggesting long-term effects.

Table 3 presents the overall estimated relative risk (summing up the contributions for the 30 days of lag, ie, A+B+C) computed for a 10-unit increase in $O_3$ compared with a reference value of 120 $\mu g/m^3$. A decomposition of the $O_3$ effect in three parts A, B, and C as in Figures 1 and 2 are also presented in Table 3. The overall RR for emergency room visits is 1.15 (95% CI: 1.01-1.29). As expected, the severity of the health effects caused by increased ozone levels is greater for individuals that require hospital admission than for those individuals looking for emergency care. The overall RR for hospital admissions is 1.42 (95% CI: 1.02-1.97). In each case, long-term effects play a major role in the overall results. In fact, short-term displacement exerts a small effect both on emergency room visits and on hospital admissions. On the other hand, short-term harvesting is substantial to both outcomes, representing the depletion of the risk pool.

Table 3. Relative risk estimates and 95% confidence intervals

<table>
<thead>
<tr>
<th>Periods</th>
<th>Emergency Room Visits</th>
<th>Hospital Admissions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RR</td>
<td>95% CI</td>
</tr>
<tr>
<td>A</td>
<td>1.04</td>
<td>0.98</td>
</tr>
<tr>
<td>B</td>
<td>0.95</td>
<td>0.92</td>
</tr>
<tr>
<td>C</td>
<td>1.16</td>
<td>1.13</td>
</tr>
<tr>
<td>A+B</td>
<td>0.99</td>
<td>0.90</td>
</tr>
<tr>
<td>A+B+C</td>
<td>1.15</td>
<td>1.01</td>
</tr>
</tbody>
</table>

It is important to notice that the sum of periods A+B underestimates the public health significance of average daily ozone concentrations. Had the analysis failed to allow for a relatively long distributed lag, no statistically significant association would have been found between ozone levels and morbidity. Period C, corresponding to the sum of the last 20 days for emergency room visits and to the sum of the last 21 days for hospital admissions, is responsible for the high relative risks observed. This may be due to an increase of the children at risk [10, 11], or to harmful effects of ozone exposure that persist for an extended time, or to both.
4. Conclusions

In the present study, we investigate the temporal pattern and the general effect of ozone exposure on children and adolescents’ respiratory morbidity expressed in terms of daily emergency room visits and admissions due to lower airways diseases in a district hospital located in a northeast rural area of Portugal in southern Europe. The effects of ozone are estimated using the recently developed distributed lag non-linear models allowing for a relatively long timescale, while controlling for weather effects, a range of other air pollutants, and long and short term patterns. The effect of ozone associated with a 10 µg/m³ increase above the European reference value on both health end-points is found distributed over an extended time, with a cumulative relative risk of 1.15 (95% CI: 1.01-1.29) for emergency room visits, and a cumulative relative risk of 1.42 (95% CI: 1.02-1.97) for hospital admissions. The public health significance of these effects is higher than has been previously reported in the literature, providing evidence contrary to the conjecture that the ozone-morbidity association is mainly due to short-term harvesting. In fact, our data analysis reveals that the effects of ozone at medium and long timescales (harvesting-resistant) are substantially larger than the effects at shorter timescales (harvesting-prone), a finding that is consistent with all children and adolescents being affected by high ground-level ozone, and not just the very frail.

References


