Improved COPD forecast algorithm incorporating atmospheric composition variables

Christophe Sarran
May 2011

1. Introduction

Initial descriptive analyses of COPD emergency hospital admissions in the United Kingdom (UK) using the EURAD (RIU) and SILAM (FMI) air quality models, as well as the Met Office NAME, found that there was a link between risk of hospitalisation for COPD and air quality, with nitric oxide concentrations being the strongest predictor. Follow-on analyses of modelled air quality indicators against measured COPD exacerbation events using generalised additive models found the concentrations of nitrogen dioxide where a significant predictor of exacerbation events. It was therefore proposed to repeat the COPD algorithm development analysis using atmospheric composition parameters as well meteorological ones.

2. Method

Following from the reviews of and recommendations from the latest COPD algorithm development (2009), a simple stepwise linear regression method was used to analyse COPD emergency hospital admission rates r against suitably transformed parameters X_i, X_i', X_i'' of meteorology and atmospheric composition.

The general linear model is:

$$\ln r = B + \underbrace{\sum_i a_i X_i}_{\text{UK}} + \underbrace{\sum_i a_i' X_i'}_{\text{PCA}} + \underbrace{\sum_i a_i'' X_i''}_{\text{PCD}} + \varepsilon$$

where ε is the residual to be minimised, the steps being first to compute coefficients a_i for the UK as a whole, next coefficients a_i' based on rates for postcode areas PCA and finally coefficients a_i'' based on rates for postcode districts PCD, with B a constant.

At each step the coefficient a is selected if it satisfies the t-test $|t| \geq t_{crit}$ for the null hypothesis h_0 with alternative hypothesis h_a such that:

$$\text{I} \quad \begin{cases} h_0 : a = 0 \\ h_a : a \neq 0 \end{cases} \qquad \text{II} \quad \begin{cases} h_0 : a = 0 \\ h_a : a \geq 0 \end{cases}$$

The 2-tailed t-test (I) is used for meteorological parameters of temperature, relative humidity and pressure, while the 1-tailed t-test (II) is used for other meteorological parameters and all parameters of atmospheric composition. For this purpose, t_{crit} is calculated to reject h_0 with a $p \leq 0.1$ probability of error using the number of weeks analysed for the degrees of freedom.

3. Result

The analysis provides a list of significant parameters that are subsequently used in the COPD algorithm to deliver the Met Office COPD forecasts. These include a lag L, a cumulative period C along with the regression coefficient a. UK results (Table 1) are computed on the strength of the measurement of hospitalisation risk over the whole of England, while PCA results (Table 2) take into account spatial variations in the hospitalisation risk with respect to spatial variations of weather and air pollution exposure. Analysis using a probability of error of $p \leq 0.05$ returned the three UK weather parameters only (and no PCA parameters).

Christophe Sarran
May 2011

Table 1. UK results.

Parameter	**Lag** *L*	**Cumulative period** *C*	**Regression coefficient** *a*
Maximum temperature	5 days	18 days	-0.93
PM_{10} particulate matter	0 days	9 days	0.050
Surface pressure	0 days	1 day	-0.0185
Precipitation	57 days	17 days	1240
Nitrogen dioxide	88 days	5 days	0.065
Ozone	33 days	32 days	0.0195

Table 2. PCA results.

Parameter	**Lag** *L*	**Cumulative period** *C*	**Regression coefficient** *a*
Maximum temperature	3 days	19 days	-1.61
Mean temperature	0 days	20 days	-3.3
Minimum temperature	0 days	18 days	-6.8
Carbon Monoxide	0 days	21 days	10.0
Relative humidity	6 days	91 days	52
Nitrogen dioxide	0 days	91 days	28
Nitric oxide	45 days	91 days	15.0
Wind speed	0 days	40 days	23
Precipitation	21 days	91 days	54×10^{4}
Sulphur dioxide	22 days	58 days	22

The UK results show that PM_{10}, NO_2 and O_3 all are predictors of risk to COPD patients. Furthermore, PCA results include CO, NO and SO_2. It can be seen that more parameters where selected at the higher spatial resolution (PCA) than at UK level. This suggests that, with even higher spatial resolution, the PCD level analysis may yield significant results with $p \leq 0.05$ especially given the argument of the localised effect of atmospheric pollution in urban areas.

4. Recommendation

Using a selection of atmospheric composition variables, it is possible to improve the COPD forecasts. It is therefore recommended to (a) complete the analysis at PCD level with $p \leq 0.05$ to confirm the selection of air quality predictors and (b) implement the algorithm described by the coefficients stated above for a trial period over the Summer of 2011. This trial would run alongside the existing operational COPD service to evaluate its performance.