Article

Estimating the Impact of Air Pollution on Healthcare-Seeking Behaviour by Applying a Difference-in-Differences Method to Syndromic Surveillance Data

Roger Morbey [1,*], Gillian Smith [1], Karen Exley [2], André Charlett [3], Daniela de Angelis [3], Sally Harcourt [1], Felipe Gonzalez [4], Iain Lake [4], Alec Dobney [2] and Alex Elliot [1]

1   Real-Time Syndromic Surveillance Team, Field Service, UK Health Security Agency, Birmingham B2 4BH, UK; gillian.smith@ukhsa.gov.uk (G.S.); sally.harcourt@ukhsa.gov.uk (S.H.); alex.elliot@ukhsa.gov.uk (A.E.)
2   Chemicals and Toxicology Science, UK Health Security Agency, London SE1 8UG, UK; karen.exley@ukhsa.gov.uk (K.E.); alec.dobney@ukhsa.gov.uk (A.D.)
3   Statistics, Modelling and Economics, Analytics and Data Science, UK Health Security Agency, London NW9 5EQ, UK; andre.charlett@ukhsa.gov.uk (A.C.); daniela.deangelis@ukhsa.gov.uk (D.d.A.)
4   School of Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, UK; felipe.gonzalez@uea.ac.uk (F.G.); i.lake@uea.ac.uk (I.L.)
*   Correspondence: roger.morbey@ukhsa.gov.uk

Abstract: Syndromic surveillance data were used to estimate the direct impact of air pollution on healthcare-seeking behaviour, between 1 April 2012 and 31 December 2017. A difference-in-differences approach was used to control for spatial and temporal variations that were not due to air pollution and a meta-analysis was conducted to combine estimates from different pollution periods. Significant increases were found in general practitioner (GP) out-of-hours consultations, including a 98% increase (2–386, 95% confidence interval) in acute bronchitis and a 16% (3–30) increase in National Health Service (NHS) 111 calls for eye problems. However, the numbers involved are small; for instance, roughly one extra acute bronchitis consultation in a local authority on a day when air quality is poor. These results provide additional information for healthcare planners on the impacts of localised poor air quality. However, further work is required to identify the separate impact of different pollutants.

Keywords: public health; epidemiology; health burden; air pollution; syndromic surveillance

1. Introduction

Air pollution is a major cause of health problems across the world [1,2]. In addition to increased mortality, air pollution can increase morbidity with a range of symptoms, including: cough and wheezing [3], childhood asthma [4,5], and cardiopulmonary [6] and ischaemic heart disease [7]. Indeed, one study in Washington DC found that increases in ozone (O_3) were associated with increased healthcare utilisation but not hospitalisations, whilst another study identified impacts on educational attainment, suggesting that monitoring just mortality or hospital admissions will underestimate the burden of impact [8,9]. Therefore, periods of poor air quality (when pollutant levels are increased) will lead to an increased demand for healthcare services [10–12].

Many countries routinely monitor air quality for the presence of pollutants and population health so that they can manage healthcare demand and take appropriate action, for example, communicating with at-risk groups [13]. The impact of infectious disease on population health is monitored through traditional laboratory surveillance and syndromic surveillance systems. Syndromic surveillance captures changes in symptom trends through monitoring electronic health records or other datasets [14]. Whilst laboratory surveillance is essential for identifying pathogens responsible for seasonal outbreaks of disease at the national, regional, and local level, syndromic surveillance can also monitor impacts on

public health caused by environmental factors, such as air pollution [15–17]. However, it is unclear whether temporary increases in healthcare demand caused by localised air pollution can be detected by existing syndromic systems.

In this study, daily air pollutant and syndromic surveillance data were collected in England, United Kingdom (UK), to estimate the impact of periods of local poor air quality on healthcare demand. Furthermore, confounding factors that would bias estimates, including public holidays or deprivation, were accounted for. By using routine monitoring data, this research has direct relevance to surveillance systems, identifying which syndromic indicators are the most useful for the routine surveillance of air pollution.

2. Method

Data on air pollution were obtained from UK Air (uk-air.defra.gov.uk, accessed on 11 May 2019) for the following air pollutants: sulphur dioxide (SO_2), nitrogen dioxide (NO_2), ozone (O_3), and particulate matter at 2.5 and 10 microns ($PM_{2.5}$ and PM_{10}). Daily data were collected between 1 April 2012 and 31 December 2017, from 96 English monitoring sites, covering 67 upper-tier local authorities (UTLAs) (England has 152 UTLAs, ranging in size from less than 100,000 to over 1 million population). These five pollutants are regularly monitored with daily forecasts to provide actions and health advice when levels are high enough to constitute a threat [18]. There are established daily air quality index (DAQI) thresholds in the UK describing low, moderate, high, or very high levels of pollutants [19]. For this study a location was defined as 'exposed' to an air pollutant if a monitoring site within the upper-tier local authority (UTLA) had recorded daily levels above the moderate threshold. By contrast a UTLA was labelled as a control if all sites in the authority had recorded daily levels below the moderate threshold with no missing data for any pollutant. The moderate DAQI threshold (break points) was taken as the boundary for exposure as this complies with existing recommendations for modifying behaviour for at-risk groups to protect from the health threat of air pollution. Therefore, these results will have direct relevance to current public health forecasting practice. A period of exposure was defined as having at least one day when pollutants were at a moderate or higher level and at least seven days prior to exposure when levels were consistently below the moderate threshold.

Syndromic data were obtained from three national Public Health England syndromic surveillance systems, including family doctor, known as general practitioners (GPs) in the UK, consultations (in hours and out of hours with unscheduled care), and calls to a National Health Service (NHS) telephone helpline (NHS 111) [20] Harcourt, 2012 #11 Harcourt, 2016 #846. Existing syndromic indicators which were routinely monitored in each system were selected based on expert clinical and epidemiological knowledge of those indicators which may be sensitive to air pollution, including: GP consultations for stroke, chest pain, eye irritation, cardiac problems, acute bronchitis, acute presenting asthma, difficulty breathing, pharyngitis, conjunctivitis, or allergic rhinitis; also, NHS 111 calls for sore throat, eye problems, difficulty breathing, and cough. In addition, as a sensitivity analysis one gastrointestinal indicator was included from each system (gastroenteritis from GP systems and 111 calls for diarrhoea), which should not be affected by air pollution.

To estimate the direct effects of air pollution on health indicators it was necessary to compare with controls where high levels of air pollution were not present. Controls can either be defined spatially, as other UTLAs during the same period that are not exposed, or as the same exposed UTLAs at different time periods (when not exposed). However, syndromic indicators may vary systematically between different locations and rates can change over time due to reasons other than air pollution. Therefore, in this study, a 'difference-in-differences' (DiD) method was applied to estimate and account for separately spatial and temporal effects and, thus, enable a direct estimate of the impact of air pollution. For each 'period of exposure' a UTLA was defined as a control if all its monitoring sites had levels below the moderate threshold, with no missing data either during the exposure period or for the 7 days before. The advantage of the DiD method is in accounting for spatial variables (e.g., deprivation) that cause differences between UTLAs which should

not vary in the short term, and temporal variables (e.g., day of the week) that are likely to have similar effects across all UTLAs. This is important as it simplifies the analysis, particularly where variables are difficult to measure.

Each of the 'periods of exposure' was treated as a separate study, with regression models used to estimate the direct impact of exposure to air pollution. Furthermore, explanatory variables included lags of up to three days to allow for delays in their impact on pollution or syndromic data. Lags were investigated using separate models to allow for the possibility of a delay between exposure to pollution and presenting with health problems. Regression was carried out in R using the MASS package [21], using a negative binomial of syndromic count data with total system activity as an offset. Equation (1) shows the general form for models:

$$y_{st} = \ln(N_{st}) + E_{st} + L_s + P_t + X_{st} + D_t \tag{1}$$

The suffix, s represents the spatial dimension (UTLA) and t the temporal dimension (date). y is the daily count within the UTLA. N is the total activity in the syndromic system. E represents exposure to air pollution and is 1 for exposed UTLAs during the period of exposure, 0 otherwise. E gives the direct estimate for impact of air pollution. L represents location, 1 for those UTLAs that experienced levels above the moderate thresholds, 0 otherwise. L provides the estimate for differences between UTLAs that do not vary over time. P is 1 during the exposure period and 0 for the 7 days prior. P represents any temporal differences that are not due to air pollution. X is a matrix of other explanatory variables, including lags of up to 3 days. D is a factor for each date, to allow for day of week and other temporal effects.

Individual periods of exposure could be as short as one day and could involve a small number of UTLAs being exposed. Low numbers of data points result in considerable uncertainty for estimates of the direct effects of air pollution. Therefore, meta-analysis was used to combine the estimates and reduce the overall uncertainty. For each syndromic indicator, the estimates for each exposure period were combined to give a joint estimate for the impact of air pollution that was more precise. Estimates for each exposure period were inversely weighted by its standard deviation, thus, more weight was given to the more precise estimates. Separate meta-estimates were created for different possible lags between exposure and seeking healthcare as whilst it was not known whether lags would be important, it was likely that the same delays would apply to each period. Meta-analysis was appropriate for this study because, although different exposure periods would involve different UTLAs as cases or controls, the methodology is the same in each case.

3. Results

To create a directed acyclic graph (DAG) for the analysis, expert opinion and published literature were elicited to list many different factors that could have an impact on either the level of air pollution in a UTLA or the incidence of healthcare-seeking behaviour. These factors included socio-economic variables (e.g., deprivation), environmental factors (e.g., weather and pollen), respiratory pathogens (e.g., influenza), and the potential impact on healthcare-seeking behaviour from media reporting of pollution forecasts. A DAG was created using these variables to describe the understanding of their causal relationship (Figure 1). Many variables were considered unlikely to change in the short term, between exposure periods and the preceding control period of 7 days, e.g., deprivation or seasonality [22]. Therefore, these variables could be accounted for by measuring the difference between case and control UTLAs, both before and during the exposure periods. Additionally, some variables were likely to vary between the exposure period and during the previous day's control period, e.g., public holidays but have similar effects across all UTLAs. Therefore, these variables could be accounted for by the estimates for the difference between the exposure period and preceding days across all UTLAs. The main remaining confounding factors potentially impacting on both pollution levels and syndromic data (and differing across time and space) were weather and pollen [23,24]. Therefore,

explanatory weather variables were included based on daily rainfall (mm), humidity (%), and temperature (°C). In addition, temperature extremes were considered an important factor and, therefore, 'temperature squared' was included to allow for a more complex relationship than a simple linear variable [22]. The potential confounding effects of pollen and spores were accounted for by being included in the PM_{10} pollutants variable.
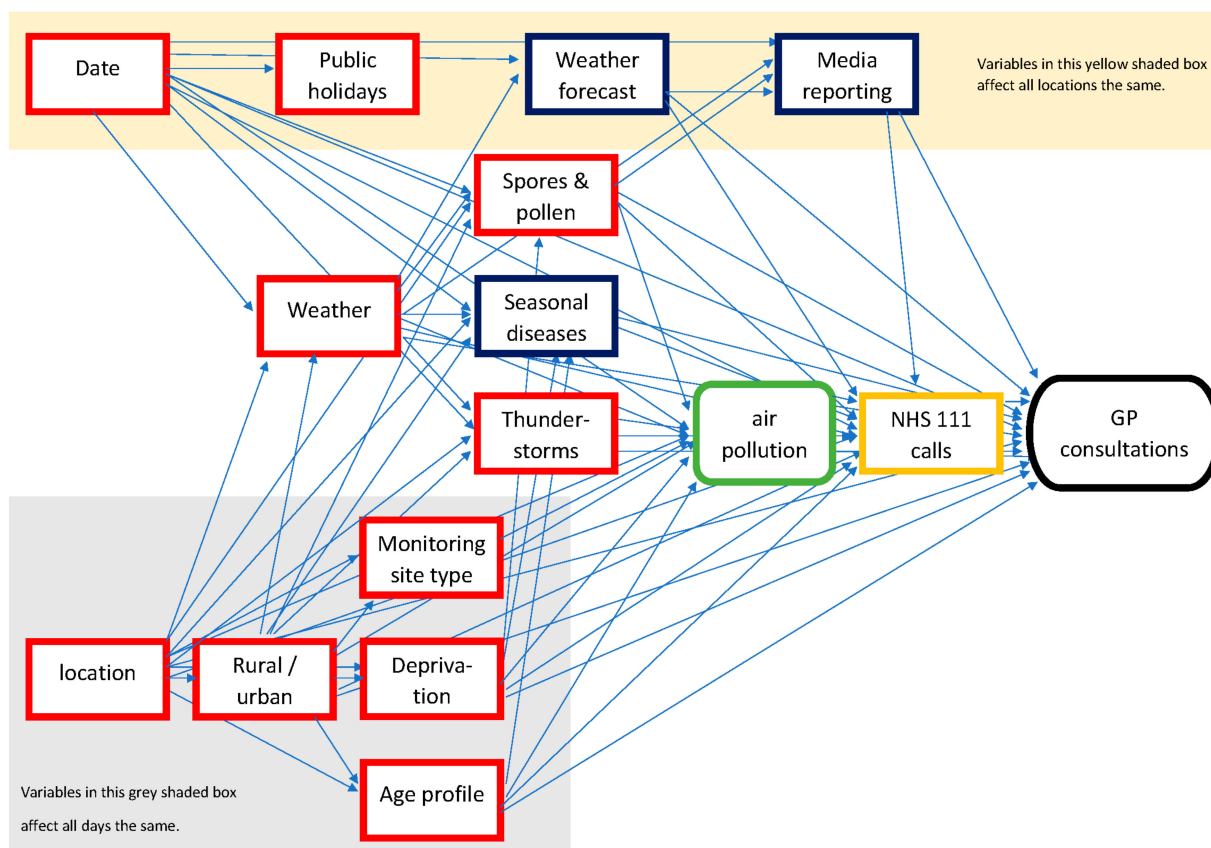

Figure 1. DAG to investigate impact of air pollution on GP consultations. Red boxes depict confounding variables that affect both the exposure of air pollution and the outcome, GP consultations. Blue boxes show variables that are competing exposures, affecting GP consultations but not pollution.

Pollutant data were available across all 2101 days included in the study, including over 193,000 site readings (Table 1). However, no readings were available where SO_2 exceeded the moderate threshold and NO_2 reached the moderate threshold of 201 µg/m^3 in only one reading (Table 1).

It was not possible to calculate separate estimates for the impact of each type of pollutant. Firstly, there were no examples of SO_2 or NO_2 pollution during the study period that were above the moderate threshold. Secondly, periods with high PM_{10} pollution were found to be coincidental with $PM_{2.5}$. PM_{10} and $PM_{2.5}$ have a strong correlation as they are fractions of the same pollutant (often the same source), so high PM_{10} readings will normally occur in parallel to high $PM_{2.5}$ (and same for the reverse). Finally, because one pollutant could act as a competing exposure when trying to estimate the impact of another pollutant [22], a UTLA was only included as a control when estimating the impact of O_3 if corresponding data, confirming that its readings for particulate pollutants were below the moderate threshold, were available. Unfortunately, eliminating the possibility of one pollutant being a confounder when measuring another meant excluding a very high proportion of the data. Figure 2 shows the daily readings for O_3, with non-grey markers showing data points that could be used to estimate the specific impact of O_3. Whilst many days are excluded because levels everywhere were below the moderate threshold, most readings above the threshold could not be included because particulate levels might also be

Int. J. Environ. Res. Public Health **2022**, *19*, 7097

5 of 16

high in the same UTLA on the same day. Therefore, air pollution exposure was defined as a day when the moderate DAQI was exceeded for any pollutant. Table A1 lists the exposure periods and the number of UTLAs available for cases and controls. Further, 78 separate exposure periods were identified with pollutant levels above moderate thresholds; however, 24 of these could not be used because of missing exposure data during the preceding 7-day control period.

Table 1. Summary of pollutant data illustrating the threshold levels and the number of days when these were exceeded during the study period (1 April 2012–31 December 2017).

	Pollutant	Daily Air Quality Index (DAQI)			Max Value Recorded
		Moderate	High	Very High	
Lower threshold for DAQI level ($\mu g/m^3$)	NO_2	201	401	601	201
	O_3	101	161	241	144
	PM_{10}	51	76	101	210
	$PM_{2.5}$	36	54	71	102
	SO_2	267	533	1065	73
		Moderate	High	Very High	All levels
Number of site readings within DAQI band	NO_2	1	0	0	193,098
	O_3	196	0	0	114,952
	PM_{10}	1109	98	7	83,888
	$PM_{2.5}$	2006	345	57	99,725
	SO_2	0	0	0	36,657
Number of unique days with a reading within DAQI band	NO_2	1	0	0	2101
	O_3	67	0	0	2101
	PM_{10}	215	32	6	2101
	$PM_{2.5}$	212	51	17	2101
	SO_2	0	0	0	2101

Pollutants: sulphur dioxide (SO_2), nitrogen dioxide (NO_2), ozone (O_3), and particulate matter at 2.5 and 10 microns ($PM_{2.5}$ and PM_{10}).

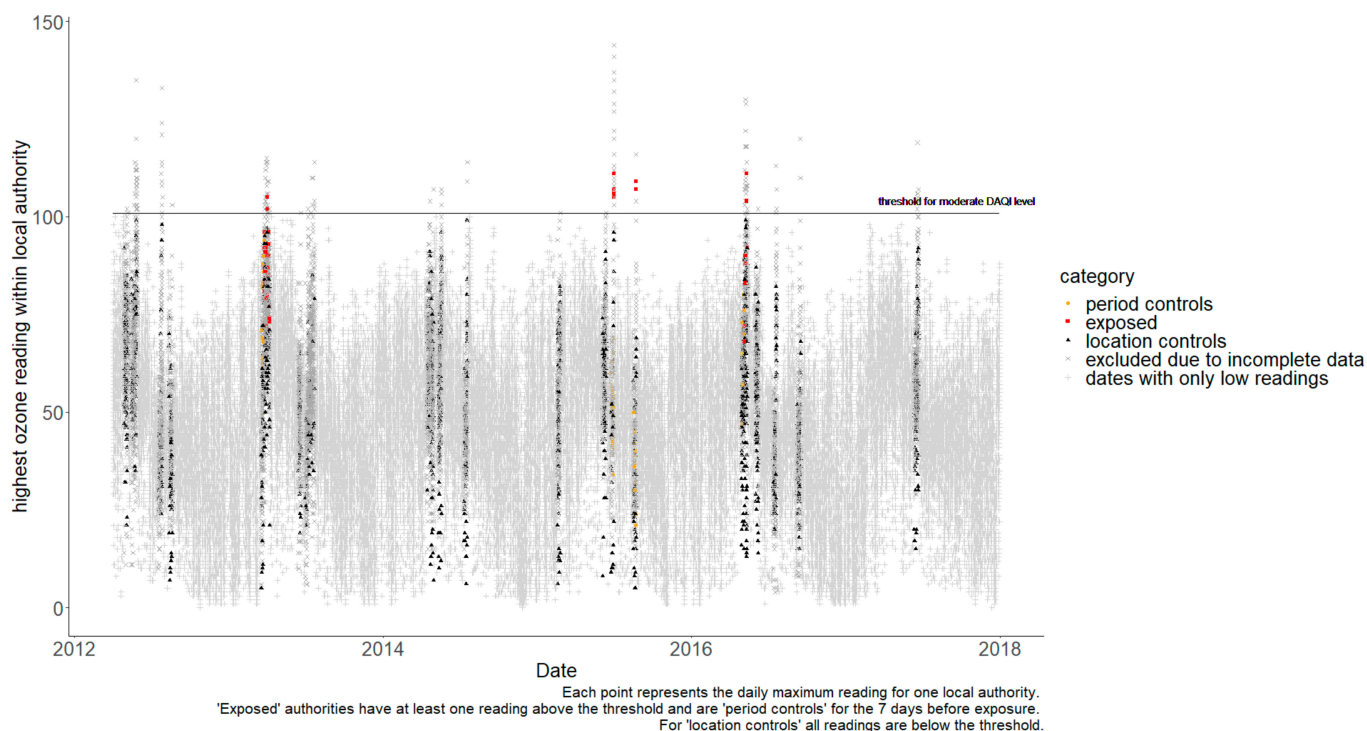**Figure 2.** Daily maximum readings for O_3 by local authority.

Syndromic data were available for every day of the study period, including a variable of 'total activity' for each surveillance system, which was used as a proxy for system coverage (as this can vary daily). Daily syndromic counts can vary considerably as underlying incidence of syndromes varies, some indicators are more specific than others, and coverage

can vary by UTLA. The smallest mean daily count was 0.2 for GP out-of-hours stroke consultations and the highest 19.4 per 1000 registered patients for GP in-hours gastroenteritis consultations (Table 2).

Table 2. Meta analysis results—estimate of rate ratios for direct effect of air pollution on syndromic indicators. Results where lower 95% confidence interval is at least 1 are highlighted in bold.

System: Syndrome	Lag (Days)	Rate Ratio	95% Confidence Interval	
GPOOH: acute bronchitis	**0**	**1.98**	**1.02**	**3.86**
GPOOH: Stroke	1	1.45	0.40	5.34
GPOOH: Eye irritation	**4**	**1.35**	**1.01**	**1.80**
GPOOH: chest pain	**3**	**1.27**	**1.01**	**1.58**
GPOOH: Cardiac	**0**	**1.26**	**1.06**	**1.49**
GPOOH: Asthma/Wheeze/DB	**2**	**1.17**	**1.00**	**1.35**
NHS 111: eye problems	**0**	**1.16**	**1.03**	**1.30**
GPOOH: Gastroenteritis	**1**	**1.13**	**1.02**	**1.25**
GPIHSS: allergic rhinitis	3	1.07	0.96	1.20
NHS 111: sore throat	1	1.07	0.97	1.19
GPIHSS: acute presenting asthma	2	1.06	0.95	1.17
NHS 111: diarrhoea	4	1.06	0.96	1.16
GPIHSS: pharyngitis or scarlet fever	3	1.04	0.98	1.10
NHS 111: cough	1	1.04	0.95	1.13
NHS 111: difficulty breathing	0	1.03	0.95	1.12
GPIHSS: gastroenteritis	4	1.02	0.98	1.07
GPIHSS: conjunctivitis	0	1.00	0.95	1.06

Syndromic Surveillance Systems: GP out-of-hours and unscheduled care (GPOOH), GP in hours (GPIHSS), National Health Service telephone helpline (NHS 111).

Estimates for the direct impact of air pollution for individual exposure periods were calculated and visualised using forest plots. There was a lot of uncertainty and the 95% confidence intervals were wide. Figure 3 shows a forest plot for the individual estimates of each exposure period on GP out-of-hours consultations for acute bronchitis (with a lag of 0 days). The meta-analysis estimate combining all periods has a much shorter confidence interval, representing the greater certainty achieved by combining data from all periods.

The meta-analysis resulted in estimates for the rate ratio (RR) of the direct effect of air pollution, which ranged from RR 0.55 (0.12–2.28) for same day GP out-of-hours (GPOOH) stroke consultations, to RR 1.98 (1.02–3.86) for same day GP out-of-hours acute bronchitis consultations. All confidence intervals in this study were calculated to be statistically significant at a 95% level. Table 2 shows the lags with the highest estimates for each indicator, with all lags available in Table A3. The highest RR was GPOOH acute bronchitis with the next highest estimate GP out-of-hours stroke consultations at a 1-day lag (RR 1.45 [CI 0.40–5.34]; however, the uncertainty around GPOOH stroke estimates was very high (Table 2). Statistically significant estimates were obtained for other GPOOH indicators, including cardiac consultations (0-day lag); acute presenting asthma, wheeze, and difficulty breathing problems (+2-day lag); chest pain (+3-day lag); and eye irritation (+4-day lag). There was also a significant association between air pollution and NHS 111 calls for eye problems at both 0- and +1-day lag. Estimates for the impact on GP in-hours consultations were smaller and none of them were statistically significant at the 95% level.

Comparisons for the meta-analysis across all syndromes are shown visually in a Forest plot, Figure 4. Whilst some of the highest effects were measured for GPOOH syndromes, the sparse data resulted in much wider confidence intervals, as depicted by smaller box sizes.

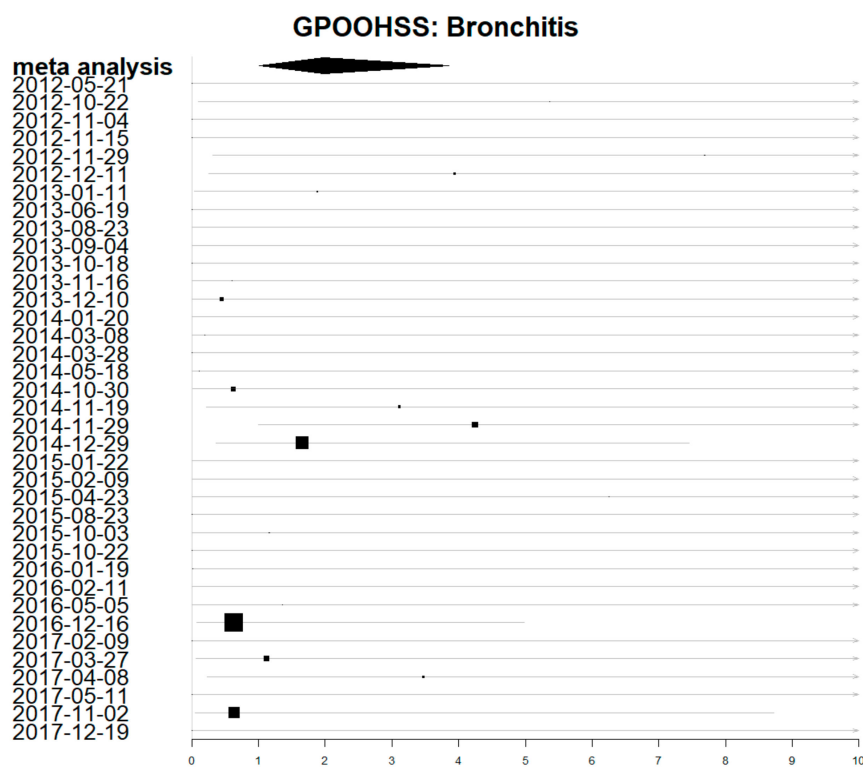**Figure 3.** Forest plot showing estimates of rate ratio for the impact of air pollution on GP out-of-hours acute bronchitis consultations. Dates are the first day of each exposure period. Larger boxes depict more precise estimates which contribute more weight.
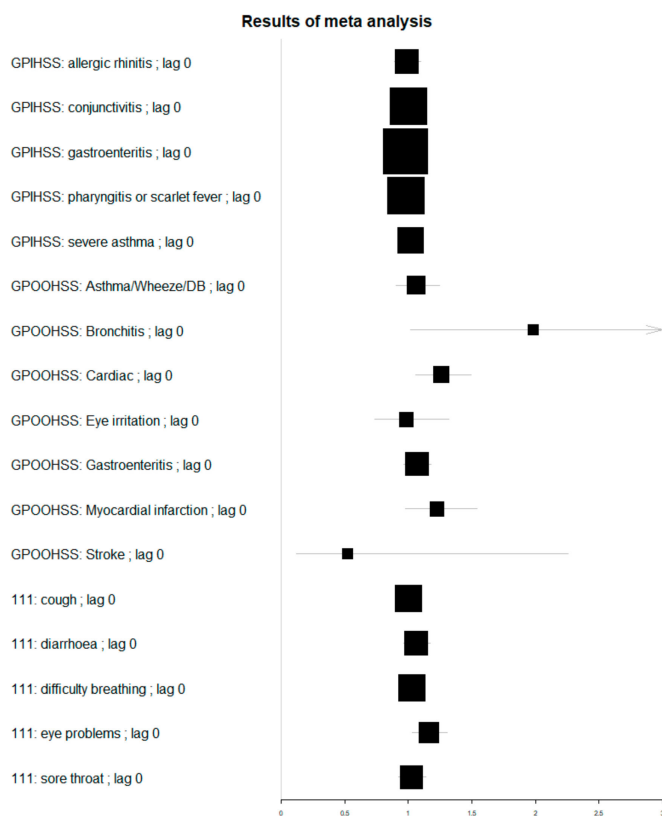

Figure 4. Forest plot of estimates for rate ratio of impact of air pollution on syndromic indicators with a zero lag. Box size is proportional to estimate precision.

4. Discussion

In this paper the direct impact of air pollution on the number of patients presenting with a range of different symptoms was estimated. Many of the symptoms included in this study (e.g., difficulty breathing and eye irritation) can have multiple aetiologies, including infectious diseases and environmental factors other than air pollution, e.g., pollen and spores. It is, therefore, important to try and disentangle the direct impact of air pollution from other confounding variables to understand its contribution. Historically, researchers built predictive models, including confounding variables, to assess the relative contribution of air pollution. However, these predictive models are limited in their interpretation for individual coefficients. For example, the impact of an increase in air pollution can only be compared to the idealised, and potentially unrealistic, scenario where all other coefficients are kept constant. Therefore, there has been a lot of interest in the development of causal inference methods to provide estimates of the direct effects of air pollution [25]. The approach used in this study was to use two causal inference techniques, DAGs and DiD.

One of the main problems in estimating the impact of poor air quality using healthcare data is the number of other variables that may also trigger similar clinical presentations or symptoms, for instance, seasonal influenza epidemics. Furthermore, seasonal and socio-economic factors will result in temporal and spatial variations in syndromic data. Therefore, DAGs [25] were used to identify variables that could bias the estimates for the impact of air pollution. Using DAGs helps to formalise the assumptions behind a study and identify the role of different variables. For instance, do variables affect both the exposure (air pollution) and the outcome (syndromic data)? If so, they are 'confounders' or just affect the outcome as 'competing exposures'. Importantly, failing to account for confounders can lead to bias in estimates for direct effects (of air pollution), whilst competing exposures can increase the uncertainty around the size of estimates. Furthermore, if there had been any variables that acted as 'mediators' between the exposure and outcome, these would have to be excluded from the analysis to avoid introducing 'collider bias' to the estimates of the direct effects of air pollution [26].

The DAG enabled consideration and identification of the causal effects of various factors, which were either stable in the short term (e.g., deprivation) or stable across location (e.g., public holidays in England). Consequently, by applying a DiD method to compare cases with both control locations and control time periods, it was possible to account for many of the confounding factors. Finally, the DAG identified the factors that were not stable across time or space (e.g., weather) and, therefore, needed to be included in regression models.

The object of public health surveillance is to provide useful information for action. Therefore, we used routinely available data, including meteorological metrics and ongoing syndromic surveillance systems. Consequently, the results can validate the usefulness of syndromic surveillance for monitoring the impact of days of poor air quality, as currently measured using the DAQI. Furthermore, the syndromic indicators give an estimate for the impact on health services of air pollution, enabling pollution forecasts to be turned into forecasts of changes in demand for health services.

The largest estimated increase in syndromic rates due to poor air quality was 1.98 (1.02, 3.86) for the GP out-of-hours acute bronchitis indicator. A typical UTLA with good coverage will record around 430 GPOOH consultations a day, of which 0.7 will have acute bronchitis and 8.2 cardiac symptoms. This central estimate for the impact of a day with poor air quality would translate to an extra 0.7 (0.0, 2.0 95% CI) acute bronchitis consultations and an extra 2.1 (0.5, 4.0 95% CI) cardiac patients. However, days with poor air quality were relatively rare in the period studied.

A range of syndromic indicators were compared across three syndromic systems. Almost all the significant increases were detected in the GP out-of-hours system; no GP in-hours syndromes giving increases were statistically significant. This could reflect differences in the case mix of patients or the availability of services. GP in-hours services require booking an appointment in advance, limiting the availability of same-day consultations,

and are only open during office hours. By contrast, GP out-of-hours services are available at weekends and evenings. It is likely, that patients use these services differently, perhaps using out-of-hours for urgent acute health needs and using in-hours for chronic disease management. Similar syndromic indicators in different systems have similar names but different underlying codes; for example, the 'asthma/wheeze/difficulty breathing' GPOOH indictor represents different conditions to the GPIHSS 'acute presenting asthma' indicator. It is unlikely, that the lack of significant results for GPIHSS indicators is due to data volume, as coverage is greater than for GPOOH. Significant increases were found for a range of GPOOH indicators, including respiratory, cardiac, and eye irritation syndromes. However, for NHS 111 calls, the only significant indicator was for eye problems.

The importance of the impact of air pollution on public health is reflected in the number and range of studies in recent years. Bae et al. reviewed research papers in South Korea into the mortality and morbidity of air pollution [27]; symptoms of morbidity included asthma, respiratory and cardiovascular hospitalization, upper and lower respiratory symptoms, low birth weight, depression, insulin resistance, allergic diseases, airway hyperresponsiveness, and new episodes of wheezing. A similar country-wide review found evidence of increased respiratory and cardiovascular disease in China associated with air pollution [1].

An important limitation in population studies, particularly involving syndromic data, is that no direct link can be shown between exposure and morbidity. There is uncertainty around whether the patients presenting with symptoms have been exposed to high levels of the pollutants recorded in their locations and, therefore, a causal link for individuals cannot be proven. Similarly, there may be very complex causal interactions, for example, between weather and pollutant variables, that have not been accounted for. Finally, there is always the possibility of other unknown or unmeasured variables, e.g., indoor pollutants that could bias results.

Whilst many of the studies mentioned in the literature reviews above provide estimates for specific pollutants, this study has not attempted to distinguish between the impacts of different pollutants. Primarily, this was due to limitations in the available data, with no examples of periods with high levels of nitrogen dioxide NO_2 or sulphur dioxide SO_2. It was not possible to disentangle the impact of PM_{10} from $PM_{2.5}$ because high levels of these pollutants usually occurred concurrently in both time and place. Finally, a conservative approach to missing data was taken, only including UTLAs as controls when there was evidence of low pollutant levels for particulate matter and O_3. Thus, there were limited examples available and insufficient data to provide separate estimates for O_3 and particulate matter.

To validate the approach, a sensitivity analysis included gastrointestinal indicators, which would not normally be affected by air pollution. GPOOH gastroenteritis consultations had an estimate that was just statistically significant at the 95% level. This was the only gastrointestinal indicator with a significant result. It suggests that results should be interpreted with some caution, particularly estimates for indicators where the increases are small and the confidence intervals are wide. However, it does not invalidate the estimates for GPOOH acute bronchitis or cardiac indicators, which were considerably larger than gastroenteritis.

The methodology was chosen to address the limitations in the study. The causal inference DiD method addressed the issues of confounders and bias due to the many factors affecting morbidity. Furthermore, this approach makes modelling assumptions explicit and transparent and provides a clearer interpretation of results than is possible with predictive modelling approaches. Similarly, the use of meta-analysis enabled more precise estimates to be obtained by combining results across different periods of poor air quality in different locations. Meta-analysis has often been criticised for combining results that are not comparable, but it is appropriate in this setting, where the same methodology and data collection is used throughout.

The impacts of air pollution across three national syndromic surveillance systems in England were studied. In future, this analysis could be repeated across the newer English

Int. J. Environ. Res. Public Health **2022**, *19*, 7097

10 of 16

syndromic systems, incorporating data from emergency departments, ambulance dispatch calls, and an online symptom checker. These systems would provide additional information on a different case-mix of patients, including those with more severe symptoms who need to attend hospital. Furthermore, the difference-in-differences approach used here to compare local areas with each other and with control periods could be extended into a more general surveillance tool to measure the impact of exposures with a wide range of causes.

5. Conclusions

The research has shown that syndromic systems are sensitive to the local impacts of air pollution, with increases seen in a range of indicators, including respiratory, cardiac, and eye irritation. Furthermore, the best indicators are likely to be from the GP out-of-hours (not GP in-hours) system. This information will have a direct impact on how syndromic surveillance is conducted and the interpretation of real-time surveillance for decision makers during air pollution episodes. Further, for some indicators, lags can be expected for up to four days, which again, are useful for interpreting daily statistical alarms following forecasts of periods of poor air quality. Finally, there is increasing interest in local authorities producing health warnings to their residents when air pollution is forecast, and it would be possible to use this approach to evaluate and monitor the impact of these warnings.

Appendix A

Table A1. Exposure periods.

Exposure Period		At Least 1 Day with Moderate or Higher DAQI Level		Low DAQI Levels Only	
Start Date	End Date	At Least 1 Day with Data	No Missing Data (Cases)	At Least 1 Day with Data	No Missing Data (Controls)
4 May 2012	6 May 2012	1	0	52	12
21 May 2012	30 May 2012	30	3	22	3
24 July 2012	27 July 2012	10	0	44	3
19 August 2012	19 August 2012	1	0	51	9
11 October 2012	11 October 2012	1	0	53	15
22 October 2012	24 October 2012	22	9	31	6
04 November 2012	06 November 2012	14	5	39	10
15 November 2012	16 November 2012	13	6	40	8
29 November 2012	02 December 2012	10	3	44	9

Number of Upper Tier Local Authorities (spanning header over the four rightmost data columns)

Table A1. *Cont.*

| Exposure Period | | Number of Upper Tier Local Authorities | | | |
| | | At Least 1 Day with Moderate or Higher DAQI Level | | Low DAQI Levels Only | |
Start Date	End Date	At Least 1 Day with Data	No Missing Data (Cases)	At Least 1 Day with Data	No Missing Data (Controls)
11 December 2012	13 December 2012	14	3	40	8
11 January 2013	25 January 2013	20	3	34	2
11 February 2013	8 March 2013	42	4	12	0
20 March 2013	12 April 2013	39	10	16	0
7 May 2013	7 May 2013	1	0	52	9
19 May 2013	19 May 2013	2	1	50	9
19 June 2013	20 June 2013	9	2	45	5
6 July 2013	22 July 2013	6	0	49	2
23 August 2013	24 August 2013	14	4	40	6
04 September 2013	05 September 2013	7	1	47	10
24 September 2013	01 October 2013	30	6	25	4
18 October 2013	18 October 2013	1	1	55	12
5 November 2013	5 November 2013	1	0	54	15
16 November 2013	28 November 2013	2	1	54	11
10 December 2013	13 December 2013	21	10	35	3
20 January 2014	21 January 2014	9	4	46	6
30 January 2014	30 January 2014	2	0	53	8
28 February 2014	28 February 2014	1	0	53	8
08 March 2014	14 March 2014	37	7	18	1
28 March 2014	04 April 2014	40	11	16	1
20 April 2014	30 April 2014	22	6	34	2
18 May 2014	20 May 2014	8	1	47	9
11 June 2014	12 June 2014	1	0	54	6
11 July 2014	21 July 2014	5	0	50	5
30 July 2014	7 August 2014	1	0	52	3
5 September 2014	29 September 2014	34	3	21	0
30 October 2014	6 November 2014	18	4	38	6
19 November 2014	21 November 2014	20	6	34	6
29 November 2014	04 December 2014	8	2	46	13
29 December 2014	31 December 2014	9	4	47	13
22 January 2015	23 January 2015	16	5	38	3
09 February 2015	15 February 2015	18	4	37	2
24 February 2015	24 February 2015	1	0	53	10
12 March 2015	20 March 2015	43	12	12	0
8 April 2015	10 April 2015	41	12	13	0
23 April 2015	24 April 2015	2	2	53	12
21 May 2015	22 May 2015	1	0	56	11
12 June 2015	02 July 2015	30	4	27	5
23 August 2015	23 August 2015	5	2	52	8
25 September 2015	25 September 2015	1	0	56	11
3 October 2015	6 October 2015	26	6	31	3
22 October 2015	5 November 2015	18	1	38	4
17 December 2015	17 December 2015	1	1	56	8
27 December 2015	02 January 2016	18	5	39	3
19 January 2016	21 January 2016	23	5	32	3
11 February 2016	11 February 2016	1	1	55	6
24 February 2016	25 February 2016	2	0	53	10

Int. J. Environ. Res. Public Health **2022**, *19*, 7097

12 of 16

Table A1. *Cont.*

| Exposure Period | | Number of Upper Tier Local Authorities | | | |
| | | At Least 1 Day with Moderate or Higher DAQI Level | | Low DAQI Levels Only | |
Start Date	End Date	At Least 1 Day with Data	No Missing Data (Cases)	At Least 1 Day with Data	No Missing Data (Controls)
10 March 2016	23 March 2016	40	10	18	0
05 May 2016	12 May 2016	23	4	35	11
07 June 2016	07 June 2016	1	0	56	12
19 July 2016	20 July 2016	3	0	53	7
08 August 2016	08 August 2016	1	0	56	6
14 September 2016	15 September 2016	7	0	50	9
25 October 2016	10 November 2016	21	7	36	4
25 November 2016	06 December 2016	32	7	25	3
16 December 2016	19 December 2016	15	5	42	9
27 December 2016	06 January 2017	10	5	46	6
18 January 2017	27 January 2017	43	9	15	0
09 February 2017	15 February 2017	40	8	19	1
27 March 2017	28 March 2017	28	10	32	8
08 April 2017	09 April 2017	6	4	54	13
30 April 2017	30 April 2017	1	0	59	14
11 May 2017	11 May 2017	4	1	55	15
18 June 2017	21 June 2017	6	0	49	11
25 September 2017	27 September 2017	11	5	45	6
16 October 2017	18 October 2017	1	0	57	12
02 November 2017	06 November 2017	18	6	42	12
19 December 2017	19 December 2017	4	2	57	17
11 December 2017	11 December 2017	1	0	59	20

Table A2. Volume of daily syndromic UTLA data 2012–2017 by syndrome.

System	Syndrome	Daily Mean Counts in UTLAs	Minimum	Maximum
GP out of hours consultations	Total with a clinical code	99.1	1	1766
	Stroke	0.2	0	10
	Acute bronchitis	0.2	0	38
	Eye irritation	0.9	0	34
	Chest pain	1.0	0	29
	Cardiac	1.7	0	38
	Asthma/Wheeze/DB	2.3	0	68
	Gastroenteritis	4.6	0	130
GP in-hours consultations	Registered patients in GP practices (thousands)	188.0	0.4	1284
	Acute presenting asthma	2.6	0	113
	Allergic rhinitis	5.8	0	463
	Conjunctivitis	8.6	0	160
	Pharyngitis or scarlet fever	15.1	0	289
	Gastroenteritis	19.4	0	287
NHS 111 telephone calls	Total calls	198.4	1	3364
	Eye problems	2.7	0	64
	Diarrhoea	4.1	0	81
	Sore throat	4.5	0	149
	Difficulty breathing	7.0	0	233
	Cough	7.0	0	467

Table A3. Meta analysis results—estimate of rate ratios for direct effect of air pollution on syndromic indicators, all lags. Results where lower 95% confidence interval is at least 1 are highlighted in bold.

System	System: Syndrome	Lag (Days)	Rate Ratio	95% Confidence Interval	
GP out of hours consultations	Asthma/Wheeze/DB	0	1.06	0.91	1.24
		1	1.09	0.93	1.27
		2	**1.17**	**1.00**	**1.35**
		3	1.05	0.91	1.22
		4	1.01	0.87	1.17
	Acute bronchitis	**0**	**1.98**	**1.02**	**3.86**
		1	1.00	0.54	1.87
		2	1.53	0.82	2.84
		3	1.11	0.58	2.12
		4	1.62	0.87	3.01
	Cardiac	**0**	**1.26**	**1.06**	**1.49**
		1	1.11	0.93	1.31
		2	1.15	0.97	1.36
		3	1.17	0.99	1.38
		4	1.09	0.92	1.30
	Eye irritation	0	0.98	0.74	1.32
		1	0.72	0.53	0.96
		2	0.83	0.62	1.10
		3	1.10	0.82	1.48
		4	**1.35**	**1.01**	**1.80**
	Gastroenteritis	0	1.07	0.96	1.18
		1	**1.13**	**1.02**	**1.25**
		2	1.08	0.98	1.19
		3	1.08	0.98	1.18
		4	1.07	0.97	1.18
	Chest pain	0	1.23	0.97	1.54
		1	1.09	0.87	1.37
		2	1.19	0.95	1.49
		3	**1.27**	**1.01**	**1.58**
		4	1.12	0.89	1.41
	Stroke	0	0.52	0.12	2.26
		1	1.45	0.40	5.34
		2	0.69	0.20	2.39
		3	0.94	0.27	3.30
		4	0.74	0.20	2.71

Int. J. Environ. Res. Public Health **2022**, *19*, 7097

14 of 16

Table A3. *Cont.*

System	System: Syndrome	Lag (Days)	Rate Ratio	95% Confidence Interval	
GP in-hours consultations	Allergic rhinitis	0	0.99	0.89	1.10
		1	1.04	0.93	1.16
		2	1.00	0.90	1.12
		3	1.07	0.96	1.20
		4	0.97	0.86	1.09
	Conjunctivitis	0	1.00	0.95	1.06
		1	1.00	0.94	1.06
		2	0.97	0.92	1.04
		3	0.98	0.93	1.05
		4	1.00	0.93	1.06
	Gastroenteritis	0	0.98	0.94	1.02
		1	0.98	0.94	1.03
		2	1.00	0.95	1.04
		3	0.97	0.93	1.01
		4	1.02	0.98	1.07
	Pharyngitis or scarlet fever	0	0.98	0.93	1.04
		1	0.98	0.92	1.03
		2	1.04	0.98	1.10
		3	1.04	0.98	1.10
		4	1.02	0.96	1.08
	Acute presenting asthma	0	1.02	0.93	1.11
		1	1.03	0.94	1.14
		2	1.06	0.95	1.17
		3	0.97	0.87	1.07
		4	0.96	0.87	1.07
NHS 111 telephone calls	Cough	0	1.00	0.92	1.09
		1	1.04	0.95	1.13
		2	1.00	0.92	1.09
		3	1.02	0.94	1.11
		4	1.01	0.93	1.10
	Diarrhoea	0	1.06	0.96	1.17
		1	1.05	0.95	1.16
		2	0.98	0.89	1.08
		3	1.03	0.94	1.14
		4	1.06	0.96	1.16
	Difficulty breathing	0	1.03	0.95	1.12
		1	1.01	0.92	1.09
		2	1.01	0.93	1.10
		3	1.00	0.92	1.08
		4	1.01	0.93	1.10

Int. J. Environ. Res. Public Health **2022**, *19*, 7097

15 of 16

Table A3. *Cont.*

System	System: Syndrome	Lag (Days)	Rate Ratio	95% Confidence Interval	
	Eye problems	**0**	**1.16**	**1.03**	**1.31**
		1	**1.16**	**1.03**	**1.30**
		2	1.09	0.97	1.23
		3	1.08	0.96	1.21
		4	1.09	0.97	1.22
	Sore throat	0	1.02	0.92	1.14
		1	1.07	0.97	1.19
		2	1.05	0.94	1.16
		3	1.02	0.92	1.13
		4	1.03	0.93	1.14

References

1. Hu, L.-W.; Lawrence, W.R.; Liu, Y.; Yang, B.-Y.; Zeng, X.-W.; Chen, W.; Dong, G.-H. Ambient Air Pollution and Morbidity in Chinese. *Adv. Exp. Med. Biol.* **2017**, *1017*, 123–151. [PubMed]
2. Samek, L. Overall human mortality and morbidity due to exposure to air pollution. *Int. J. Occup. Med. Environ. Health* **2016**, *29*, 417–426. [CrossRef] [PubMed]
3. Brabin, B.; Smith, M.; Milligan, P.; Benjamin, C.; Dunne, E.; Pearson, M. Respiratory morbidity in Merseyside schoolchildren exposed to coal dust and air pollution. *Arch. Dis. Child.* **1994**, *70*, 305–312. [CrossRef] [PubMed]
4. Hansel, N.N.; Romero, K.M.; Pollard, S.L.; Bose, S.; Psoter, K.J.; Underhill, L.J.; Johnson, C.; Williams, D.; Curriero, F.C.; Breysse, P.; et al. Ambient Air Pollution and Variation in Multiple Domains of Asthma Morbidity among Peruvian Children. *Ann. Am. Thorac. Soc.* **2019**, *16*, 348–355. [PubMed]
5. Burbank, A.J.; Peden, D. Assessing the impact of air pollution on childhood asthma morbidity: How, when, and what to do. *Curr. Opin. Allergy Clin. Immunol.* **2018**, *18*, 124–131. [CrossRef] [PubMed]
6. Nayebare, S.R.; Aburizaiza, O.S.; Siddique, A.; Carpenter, D.O.; Zeb, J.; Aburizaiza, A.J.; Pantea, C.; Hussain, M.M.; Khwaja, H.A. Association of fine particulate air pollution with cardiopulmonary morbidity in Western Coast of Saudi Arabia. *Saudi Med. J.* **2017**, *38*, 905–912. [CrossRef]
7. Xie, W.; Li, G.; Zhao, D.; Xie, X.; Wei, Z.; Wang, W.; Wang, M.; Li, G.; Liu, W.; Sun, J.; et al. Relationship between fine particulate air pollution and ischaemic heart disease morbidity and mortality. *Heart* **2015**, *101*, 257–263. [CrossRef]
8. Fuller, C.H.; Jones, J.; Roblin, D.W. Evaluating changes in ambient ozone and respiratory-related healthcare utilization in the Washington, DC metropolitan area. *Environ. Res.* **2020**, *186*, 109603. [CrossRef]
9. Mizen, A.; Lyons, J.; Milojevic, A.; Doherty, R.; Wilkinson, P.; Carruthers, D.; Akbari, A.; Lake, I.; Davies, G.A.; Al Sallakh, M.; et al. Impact of air pollution on educational attainment for respiratory health treated students: A cross sectional data linkage study. *Health Place* **2020**, *63*, 102355. [CrossRef]
10. Sarnat, J.; Sarnat, S.E.; Flanders, W.D.; Chang, H.H.; Mulholland, J.; Baxter, L.; Isakov, V.; Özkaynak, H. Spatiotemporally resolved air exchange rate as a modifier of acute air pollution-related morbidity in Atlanta. *J. Expo. Sci. Environ. Epidemiol.* **2013**, *23*, 606–615. [CrossRef]
11. Kelly, F.J.; Fuller, G.W.; Walton, H.; Fussell, J.C. Monitoring air pollution: Use of early warning systems for public health. *Respirology* **2011**, *17*, 7–19. [CrossRef] [PubMed]
12. Masselot, P.; Chebana, F.; Lavigne, É.; Campagna, C.; Gosselin, P.; Ouarda, T.B.M.J. Toward an Improved Air Pollution Warning System in Quebec. *Int. J. Environ. Res. Public Health* **2019**, *16*, 2095. [CrossRef] [PubMed]
13. Glenn, T.A., Jr. Regional air pollution warning system. *J. Air Pollut. Control Assoc.* **1966**, *16*, 22–24. [CrossRef] [PubMed]
14. Triple, S. Assessment of syndromic surveillance in Europe. *Lancet* **2011**, *378*, 1833–1834. [CrossRef]
15. Hughes, H.E.; Morbey, R.; Fouillet, A.; Caserio-Schönemann, C.; Dobney, A.; Hughes, T.; Smith, G.E.; Elliot, A.J. Retrospective observational study of emergency department syndromic surveillance data during air pollution episodes across London and Paris in 2014. *BMJ Open* **2018**, *8*, e018732. [CrossRef]
16. Smith, G.E.; Bawa, Z.; Macklin, Y.; Morbey, R.; Dobney, A.; Vardoulakis, S.; Elliot, A.J. Using real-time syndromic surveillance systems to help explore the acute impact of the air pollution incident of March/April 2014 in England. *Environ. Res.* **2015**, *136*, 500–504. [CrossRef]
17. Todkill, D.; Gonzalez, F.D.J.C.; Morbey, R.; Charlett, A.; Hajat, S.; Kovats, S.; Osborne, N.J.; McInnes, R.; Vardoulakis, S.; Exley, K.; et al. Environmental factors associated with general practitioner consultations for allergic rhinitis in London, England: A retrospective time series analysis. *BMJ Open* **2020**, *10*, e036724. [CrossRef]

18. Department for Environment, Food & Rural Affairs. UK AIR-Air Information Resource. 2021. Available online: https://uk-air.defra.gov.uk (accessed on 1 November 2021).

19. Department for Environment, Food & Rural Affairs. What Is the Daily Air Quality Index? 2021. Available online: https://uk-air.defra.gov.uk/air-pollution/daqi?view=more-info (accessed on 5 November 2019).

20. Harcourt, S.E.; Smith, G.E.; Elliot, A.J.; Pebody, R.; Charlett, A.; Ibbotson, S.; Regan, M.; Hippisley-Cox, J. Use of a large general practice syndromic surveillance system to monitor the progress of the influenza A(H1N1) pandemic 2009 in the UK. *Epidemiol. Infect.* **2011**, *140*, 100–105. [CrossRef]

21. Venables, W.N.; Ripley, B.D. *Modern Applied Statistics with S*, 4th ed.; Springer: Berlin/Heidelberg, Germany, 2002.

22. Huang, G.; Brown, P.E.; Fu, S.H.; Shin, H.H. Daily mortality/morbidity and air quality: Using multivariate time series with seasonally varying covariances. *J. R. Stat. Soc. Ser. C* **2021**, *71*, 148–174. [CrossRef]

23. Misailidou, M.; Pitsavos, C.; Panagiotakos, D.B.; Chrysohoou, C.; Stefanadis, C. Short-term effects of atmospheric temperature and humidity on morbidity from acute coronary syndromes in free of air pollution rural Greece. *Eur. J. Cardiovasc. Prev. Rehabil.* **2006**, *13*, 846–848. [CrossRef]

24. Welty, L.J.; Zeger, S.L. Are the acute effects of particulate matter on mortality in the National Morbidity, Mortality, and Air Pollution Study the result of inadequate control for weather and season? A sensitivity analysis using flexible distributed lag models. *Am. J. Epidemiol.* **2005**, *162*, 80–88. [CrossRef] [PubMed]

25. Bind, M.-A. Causal Modeling in Environmental Health. *Annu. Rev. Public Health* **2019**, *40*, 23–43. [CrossRef] [PubMed]

26. VanderWeele, T.J.; Staudt, N. Causal diagrams for empirical legal research: A methodology for identifying causation, avoiding bias and interpreting results. *Law Probab. Risk* **2011**, *10*, 329–354. [CrossRef] [PubMed]

27. Bae, S.; Kwon, H.-J. Current State of Research on the Risk of Morbidity and Mortality Associated with Air Pollution in Korea. *Yonsei Med. J.* **2019**, *60*, 243–256. [CrossRef] [PubMed]