Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# Lagged environmental effects on multi-pathogen respiratory positivity: a distributed-lag nonlinear analysis of seven Sichuan sentinel hospitals, China

**Fengmiao Hu**

453516482@qq.com

Sichuan Center for Disease Control and Prevention

**Xingyu Zhou**

Sichuan Center for Disease Control and Prevention

**Lijun Zhou**

Sichuan Center for Disease Control and Prevention

**Zhirui Li**

Sichuan Center for Disease Control and Prevention

**Shuang Dong**

Sichuan Center for Disease Control and Prevention

**Biao Yang**

Sichuan Center for Disease Control and Prevention

**Chongkun Xiao**

Sichuan Center for Disease Control and Prevention

**Article**

**Keywords:** acute respiratory infection, distributed lag non-linear model, meteorological factors, air pollution, multi-pathogen surveillance, Sichuan Province

**Posted Date:** November 13th, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-11513107/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

**Additional Declarations:** No competing interests reported.

# Lagged environmental effects on multi-pathogen respiratory positivity: a distributed-lag nonlinear analysis of seven Sichuan sentinel hospitals, China

Fengmiao Hu, MD*; Xingyu Zhou, MD; Lijun Zhou, MD; Zhirui Li, MD; Shuang Dong, MD; Biao Yang, MD; Chongkun Xiao, MD

Sichuan Center for Disease Control and Prevention, Chengdu 610041, China

*Corresponding author: Fengmiao Hu, E-mail: 453516482@qq.com

**Abstract**

Acute respiratory infections (ARI) impose a large global burden and show marked seasonality that implicates environmental drivers, yet the lagged effects of these exposures on pathogen-specific positivity remain poorly characterised. Using case-level data from seven sentinel hospitals in Sichuan, China (2024–2026) aggregated to a weekly panel of 23 014 ARI cases, we applied quasi-Poisson distributed-lag non-linear models (cross-basis, 6-week lag; trend spline; city effects; log offset) to estimate cumulative relative risks (90th versus 10th percentile) for temperature, relative humidity and six air pollutants, with within-pathogen false-discovery-rate control. At least one pathogen was detected in 51.9% of cases and co-detection in 18.2%; Haemophilus influenzae and influenza predominated. Relative humidity gave the most stable single-exposure association with respiratory syncytial virus positivity (*RR* 2.01, *95% CI* 1.06–3.82); it failed FDR control (adjusted $P = 0.130$) and was nullified by temperature co-adjustment (*RR* 1.05). H. influenzae rose with temperature and ozone and fell with PM2.5; SARS-CoV-2 tracked carbon monoxide. Twelve of 40 estimates survived FDR control. The humidity–RSV signal is robust to model specification yet fragile under co-varying exposures — the cautionary lesson — and the reusable output is a transparent screening framework that exposes such fragility, serving exposure science and environmental epidemiology.

Keywords: acute respiratory infection; distributed lag non-linear model; meteorological factors; air pollution; multi-pathogen surveillance; Sichuan Province

## 1 Introduction

Acute respiratory infections (ARI) remain a leading cause of infectious illness and death worldwide[1]. In 2019, lower respiratory infections accounted for approximately 2.5 million deaths globally and remained a leading infectious cause of death in children under five years of age[1]. The aetiological spectrum is highly heterogeneous, encompassing respiratory viruses, bacteria and atypical agents, and co-detection of multiple pathogens is common[2, 3]. This complexity complicates both clinical management and the interpretation of routine surveillance data, and underlies the growing demand which is why surveillance increasingly needs pathogen-specific, rather than syndromic, risk assessment.

In temperate and subtropical regions, ARI exhibits a reproducible seasonal pattern with peaks generally coinciding with cold and dry months, indicating that meteorological conditions are important external drivers of transmission and host susceptibility[4, 5, 6].

Ambient temperature and relative humidity influence pathogen survival, aerosol stability and mucosal barrier function, while air pollutants can impair innate immune responses and facilitate respiratory viral and bacterial infection[7, 8, 9].

The distributed lag non-linear model (DLNM) introduced by Gasparrini and colleagues[10] provides a unified framework that simultaneously describes the non-linear exposure-response relationship and the temporal distribution of effects over a defined lag window[11]. Applying DLNM to ARI surveillance is particularly

relevant because respiratory pathogens act over lagged windows spanning several weeks rather than within a single day, and because single-day exposure contrasts may miss cumulative effects[12].

Most of this evidence, however, comes from single-pathogen or single-pollutant studies — influenza or RSV against temperature, or one pollutant against a syndromic case count — so we still know little about how bacterial and viral pathogens differ in their environmental sensitivities within one surveillance system[13, 14, 15]. Children under five, who carry the heaviest RSV burden, are also the group most exposed to weather-driven swings, yet have rarely been examined on their own[4].

Within China, distributed lag non-linear models have been applied to sentinel influenza-like illness series in several cities, where low temperature was associated with increased ILI risk after adjustment for long-term trend and holiday effects[16, 17]. These studies used national influenza sentinel data and a single meteorological exposure, and did not resolve pathogen-specific effects.

The environmental-epidemiology literature on respiratory infection is overwhelmingly single-pathogen and single-exposure, which limits the comparison of bacterial and viral sensitivities within one surveillance framework[7, 8, 13, 14, 15]. Where multi-pathogen or multi-exposure analyses do exist, they tend to lack the distributed-lag structure needed for cumulative weekly effects or to omit multiplicity control across the many pathogen–exposure contrasts that such a design generates[18, 19]. Most importantly, single-exposure models routinely report associations without testing whether they survive co-adjustment for the correlated meteorological and pollution exposures — a threat that is acute precisely because these exposures are strongly inter-correlated. The present study is built around that gap. We apply non-linear distributed-lag models to a multi-centre sentinel panel linking lagged weather and air-quality drivers to the positivity of several bacterial and viral pathogens, and we embed the analysis in a pre-specified screening sequence: a priori pathogen eligibility, within-pathogen false-discovery-rate control across 40 exposure–pathogen estimates, specification sensitivity, and dual-exposure boundary testing. The reproducible methodological template, not any individual association, is the contribution we intend to generalise.

Sichuan Province offers a useful setting for this question. The province spans the Sichuan Basin and the western plateau, generating a substantial gradient in temperature, humidity and air quality across its seven national sentinel hospitals, and it operates a multi-pathogen surveillance network reporting 21 pathogen categories[20, 21, 22]. We therefore aimed to quantify the lagged associations of weekly mean temperature, relative humidity and six air pollutants with the positivity of major respiratory pathogens in ARI cases, and to compare effect patterns across pathogens and age groups. We did this with a retrospective, sentinel-based ecological time-series analysis of case-level surveillance data from 1 October 2024 to 15 October 2026.

## 2 Materials and methods

### 2.1 Study design and data sources

This retrospective, sentinel-based ecological time-series analysis is reported in accordance with the STROBE statement[23]. Case-level acute respiratory infection (ARI) surveillance records were obtained from the multi-pathogen module of the national acute respiratory infectious disease sentinel surveillance system, which operates through the Chinese Center for Disease Control and Prevention and its provincial counterpart. The study period was 1 October 2024 to 15 October 2026 and covered seven national sentinel hospitals in Chengdu, Mianyang, Zigong, Panzhihua, Nanchong, Ya'an and Dazhou, where influenza-like illness (ILI) and severe acute respiratory infection (SARI) cases were sampled according to the national technical protocol for acute respiratory infection sentinel surveillance (trial version)[24]. The protocol was formally issued on 4 December 2024; before that date the surveillance was already operated on a trial basis with identical case definitions and sampling and testing procedures, so no protocol-related change in case definition or sampling practice occurred

within the study window. The surveillance platform and its earlier analyses have been described separately[20, 21, 22]. Cases were eligible when a complete multiplex panel result covering all 21 reported pathogen categories was recorded and the specimen collection date fell within the study period. The final analytical sample comprised 23 014 cases (15 381 ILI and 7 633 SARI), aggregated to a city-week panel of 742 observations (7 cities × 106 weeks). For the distributed-lag models, the first six weeks per city were excluded to ensure a complete lag 0–6 exposure history, leaving 700 city-week observations for model fitting; the descriptive lag-correlation analysis (Table 4) used the full 742-observation panel.

Outcome definition. Throughout this paper, pathogen positivity denotes the proportion of respiratory specimens collected at these sentinel hospitals that tested positive for a given pathogen. It is a test-conditional measure among ARI cases who presented to a sentinel hospital and were sampled, not a population infection rate. Care-seeking behaviour, clinicians' sampling decisions, panel availability and testing capacity all shape which cases enter the denominator, so positivity can differ from true community infection frequency. Every relative risk below should be read against that caveat.

External exposure data were retrieved from the Open-Meteo historical API. Meteorological variables (daily mean, maximum and minimum 2-m temperature, mean and minimum 2-m relative humidity, wind speed and precipitation) were derived from the ERA5 global reanalysis[25], which provides hourly fields at 0.25° (approximately 25 km) horizontal resolution from 1940 onwards. Air pollutant concentrations (PM2.5, PM10, sulphur dioxide, nitrogen dioxide, carbon monoxide and ozone) were derived from the Copernicus Atmosphere Monitoring Service (CAMS) global atmospheric composition forecast and reanalysis datasets[26, 27], which provide 3-hourly surface fields at 0.4° (approximately 45 km) horizontal resolution. Hourly pollutant values were aggregated to daily means and subsequently to weekly means. All series were extracted at the geographic coordinates of each sentinel city (30.1°–31.5° N, 103.0°–107.5° E).

The ERA5 and CAMS products are global reanalysis datasets rather than station observations, used because no publicly accessible city-level station records cover the full study window across all seven cities; reanalysis data are widely used in environmental epidemiology where station networks are sparse[28]. Two distinct sources of exposure error must be separated. The first, grid-placement error, is the difference between the value at the chosen grid point and nearby points within the same cell; re-extracting every series at six ±0.2-degree offsets showed the two ERA5 meteorological fields to be highly insensitive to grid placement (temperature median $r = 0.999$, relative humidity 0.972) whereas the CAMS pollutant fields were not (PM2.5 minimum $r = 0.655$, offset bias up to 57% of the mean). The second, and larger, source is the systematic difference between the reanalysis product and a ground monitor at the city; station records for the study window are not publicly available, so this cannot be quantified here. Validations of the same products in China show reanalysis fields track station series reasonably in temporal correlation but depart from measured magnitudes, most for pollutants and during high-concentration episodes. Such misclassification biases associations toward the null and is why the pollutant results are reported as exploratory while the humidity result, drawn from the better-performing ERA5 fields, is weighted more heavily.[29, 30, 31].

## 2.2 Pathogen detection

Nucleic acid was extracted from each specimen and amplified by multiplex real-time fluorescence quantitative PCR using commercially available pre-mixed nucleic acid detection assays for respiratory pathogens. The same assays, extraction procedure and amplification protocol were used at all seven sentinel hospitals, and all raw amplification data were re-analysed by the provincial reference laboratory[32]. The panel was reported as 21 pathogen categories: 10 viruses, 6 bacteria, 2 fungi and 3 Mycoplasma/Chlamydia species. Influenza A and Influenza B are reported as a single "influenza virus" category throughout, and the human coronaviruses detected in this panel are likewise reported together as a single category, so the number of reported categories is

smaller than the number of assay targets. Each run included positive and negative controls and an endogenous internal control. Following the manufacturer criteria, a target was called positive when the amplification curve showed a typical S shape with cycle threshold ≤ 38 and negative when the cycle threshold exceeded 40 or no curve was detected; equivocal specimens were repeated and called positive only if the repeat result agreed. Table 1 lists all 21 reportable categories, including the four with positivity below 0.3%, so that the full panel remains visible.

### 2.3 Outcome definition and pathogen selection

The outcome was the weekly count of positive detections for each pathogen, with the weekly number of specimens tested for that pathogen as the denominator. Because several targets can be positive simultaneously, outcomes were not mutually exclusive; co-detection was defined as two or more pathogens detected in the same specimen. Because the counts of different pathogens within a specimen are not independent, we verified that ignoring this dependence does not create the artefact it could: the primary model was refitted with the numerator restricted to specimens in which the pathogen was the sole organism detected and again with the numerator restricted to specimens in which it co-occurred, keeping every city-week and all model terms fixed (Supporting Information, Table S9). Five pathogens were carried into DLNM: rhinovirus, S. pneumoniae, H. influenzae and SARS-CoV-2 met both pre-specified criteria (at least 1 000 positive cases in the analytic panel and quasi-Poisson dispersion below 2.0). RSV was included as a pre-specified exception because it is the single most important paediatric respiratory pathogen; its 861 positive cases in the analytic panel fell short of the case criterion, so its estimates should be read with lower precision and stability than those of the four fully eligible pathogens, and the Benjamini-Hochberg adjusted *P* value (0.130) should be taken as the primary basis for judging the humidity-RSV association rather than its nominal significance. All RSV estimates are reported as exploratory throughout. Influenza virus and enterovirus were excluded, the former because its weekly counts were strongly over-dispersed (dispersion 5.12) and the latter because it contributed only 758 positive cases in the analytic panel (801 across the full 742-observation panel), below the pre-specified minimum. Both are nevertheless described descriptively.

Missing data. Every eligible specimen had a complete panel result for all 21 reported pathogen categories, so no pathogen outcome was missing. The reanalysis exposure series contained no missing daily values after aggregation to weekly means. All analyses therefore used complete cases and no imputation was performed.

### 2.4 Statistical analysis

Quasi-Poisson distributed lag non-linear models[10] were fitted using the cross-basis function of the R package dlnm version 2.4.10[33]. The cross-basis was specified with a natural cubic spline for the exposure dimension (*df = 3*) and a natural cubic spline for the lag dimension (*df = 3*), with a maximum lag of 6 weeks. This horizon is biologically motivated: incubation periods span days to about two weeks and serial intervals are on the order of several days, so a six-week window encompasses incubation plus one to two transmission cycles and gives weather-conditioned changes in susceptibility and exposure time to accumulate into detectable positivity. The window was fixed a priori on these biological grounds rather than by a data-driven search, which would risk overfitting the lag structure, and with approximately 100 usable weeks per city longer horizons leave too few observations for stable estimation of the lag basis. Long-term trend and seasonality were controlled by a natural cubic spline of the continuous week index (*df = 8,* capped at 10 in the sensitivity analysis to avoid overfitting; the range 4–10 that we report reflects this constraint rather than a wider search), and sentinel city was included as a categorical covariate. Because meteorological variables and pollutants were strongly correlated, each exposure was modelled separately (single-exposure models) to avoid collinearity; consequently the independent effect of each pollutant could not be estimated, and an apparent association for a given pollutant may partly

reflect co-varying exposures or seasonality[7]. This choice does not imply that a single environmental factor operates in isolation (meteorological variables and pollutants act jointly and are themselves interdependent), only that the available sample could not support stable multi-exposure estimation of their separate contributions. Temperature associations were designated exploratory a priori because of collinearity between temperature and seasonality, and were evaluated across a range of trend-spline degrees of freedom.

Cumulative relative risks (*RR*) and 95% confidence intervals (*CI*) were computed on the log scale using the linear predictor identity: for a substitution $\Delta$ of the cross-basis row relative to a reference series, the cumulative *RR* is $\exp(\Delta'\beta)$ with variance $\Delta'V\Delta$, where $\beta$ is the coefficient vector and V the covariance matrix. This formulation guarantees that *RR* equals 1 when the substituted series is identical to the reference. The reference series was a constant exposure equal to the median observed weekly value, and exposures were contrasted at the 90th and 10th percentiles of their city-week distributions[11, 34]. Internal knots for the exposure basis were placed at the 50th percentile of the exposure distribution, and the lag basis carries a single interior knot at week 3 with boundary knots at weeks 0 and 6. Because lag-basis flexibility is an analyst choice, the primary associations were re-estimated under five alternative lag parametrisations, including knots at weeks 2 and 4, knots at weeks 1, 3 and 5, a more flexible five-knot basis and a linear lag basis, holding the exposure basis and trend spline fixed (Supporting Information, Table S6); the cumulative estimates were unchanged in direction and close in magnitude across all six.

Because a single elevated week and a uniformly elevated window carry equal weight in the cross-basis projection at $df = 3$, cumulative associations are reported over increasing windows (lag 0–1 through lag 0–6 weeks) rather than as week-specific lag coefficients. Age-stratified models were fitted for children under 5 years, ages 5–18 years, 19–64 years and 65 years or older, and group comparisons of positivity were performed with the $\chi^2$ test. For the four primary pathogen–exposure pairs, the cumulative exposure–lag–response surface (the *RR* associated with an exposure sustained over the leading 0–L lag weeks, with the remaining lags held at the median reference), was computed with the same identity and is displayed as three-dimensional surfaces and contour plots. Analyses were conducted in R 4.6.1 and Python 3.13.

Multiplicity was addressed in two ways. First, because each pathogen was assessed against eight exposures, Benjamini–Hochberg false discovery rate (FDR) control at the 5% level was applied within each pathogen, and both nominal and FDR-adjusted *P* values are reported. Second, the primary inference was restricted to the five pathogens retained a priori, and temperature associations were designated exploratory in view of their sensitivity to trend-spline flexibility. Model adequacy was assessed through the quasi-Poisson dispersion parameter, lag-1 to lag-4 autocorrelation of Pearson residuals, and the ratio of observed to expected positive counts; no formal adjustment for residual autocorrelation was made beyond the trend and seasonal spline (limitations)[34].

All methods were carried out in accordance with relevant guidelines and regulations.

## 3 Results

### 3.1 Pathogen distribution and seasonality

Among 23 014 ARI cases, at least one pathogen was detected in 11 940 (51.88%) and two or more in 4 178 (18.15%), with up to seven pathogens detected in a single specimen. Positivity exceeded 0.3% for 17 of the 21 reported pathogen categories. The three most prevalent were H. influenzae (3 680/23 014; 15.99%), S. pneumoniae (2 579/23 014; 11.21%) and influenza virus (2 304/23 014; 10.01%), followed by rhinovirus (2 100/23 014; 9.12%) and SARS-CoV-2 (1 087/23 014; 4.72%). RSV was detected in 913 cases (3.97%) and enterovirus in 801 (3.48%) (Table 1). All proportions reported below are positivity among specimens tested at these sentinel hospitals and must not be interpreted as population infection rates.

Stratified by surveillance stream, positivity differed between outpatient ILI and inpatient SARI for every pathogen examined (chi-square, all $P < 0.001$; Table 2). Influenza virus positivity was higher in ILI (12.23%) than in SARI (5.54%), whereas RSV positivity was higher in SARI (5.24%) than in ILI (3.34%); the largest difference was for SARS-CoV-2 (5.99% vs 2.16%). These differences reflect the differing case definitions, severity and age composition of the two streams rather than a single population infection rate. All major pathogens except rhinovirus showed clear winter peaks; rhinovirus peaked in autumn and spring (Figure 2c).

| Pathogen | Positive, *n* | Tested, *n* | Positivity, % |
|---|---|---|---|
| H. influenzae | 3680 | 23014 | 15.99 |
| S. pneumoniae | 2579 | 23014 | 11.21 |
| Influenza virus | 2304 | 23014 | 10.01 |
| Rhinovirus | 2100 | 23014 | 9.12 |
| SARS-CoV-2 | 1087 | 23014 | 4.72 |
| RSV | 913 | 23014 | 3.97 |
| Enterovirus | 801 | 23014 | 3.48 |
| Parainfluenza virus | 772 | 23014 | 3.35 |
| K. pneumoniae | 738 | 23014 | 3.21 |
| Adenovirus | 641 | 23014 | 2.79 |
| Human metapneumovirus | 564 | 23014 | 2.45 |
| Common coronavirus | 518 | 23014 | 2.25 |
| Group A Streptococcus | 232 | 23014 | 1.01 |
| Bocavirus | 227 | 23014 | 0.99 |
| M. pneumoniae | 197 | 23014 | 0.86 |
| Aspergillus | 80 | 23014 | 0.35 |
| C. pneumoniae | 69 | 23014 | 0.30 |
| B. pertussis | 34 | 23014 | 0.15 |
| Cryptococcus | 13 | 23014 | 0.06 |
| C. psittaci | 12 | 23014 | 0.05 |
| Legionella | 7 | 23014 | 0.03 |
| Total (21 categories) | 17568 | — | — |

Table 1 Positivity of all 21 pathogen categories in 23 014 ARI cases from seven sentinel hospitals, Sichuan Province, China, 2024–2026. Influenza A and influenza B are reported as a single influenza virus category and the human coronaviruses detected in this panel are pooled into one category, so the number of reported categories is smaller than the number of assay targets. Categories with positivity below 0.3% (Bordetella pertussis, Cryptococcus, Chlamydia psittaci and Legionella) are listed for completeness. The total row sums positives across the 21 categories and therefore exceeds the number of cases with any pathogen detected, because co-detection is common.

| Pathogen | ILI positive, *n* | ILI positivity, % | SARI positive, *n* | SARI positivity, % |
|---|---|---|---|---|
| H. influenzae | 2741 | 17.82 | 939 | 12.30 |
| Influenza virus | 1881 | 12.23 | 423 | 5.54 |
| S. pneumoniae | 1788 | 11.62 | 791 | 10.36 |
| Rhinovirus | 1495 | 9.72 | 605 | 7.93 |
| SARS-CoV-2 | 922 | 5.99 | 165 | 2.16 |
| Enterovirus | 589 | 3.83 | 212 | 2.78 |
| RSV | 513 | 3.34 | 400 | 5.24 |

Table 2 Positivity by surveillance stream: 15 381 outpatient ILI and 7 633 inpatient SARI specimens. Differences between streams were significant for every pathogen (chi-square, all $P < 0.001$). No correction for multiplicity was applied, so the nominal significance should not be read as confirmatory.

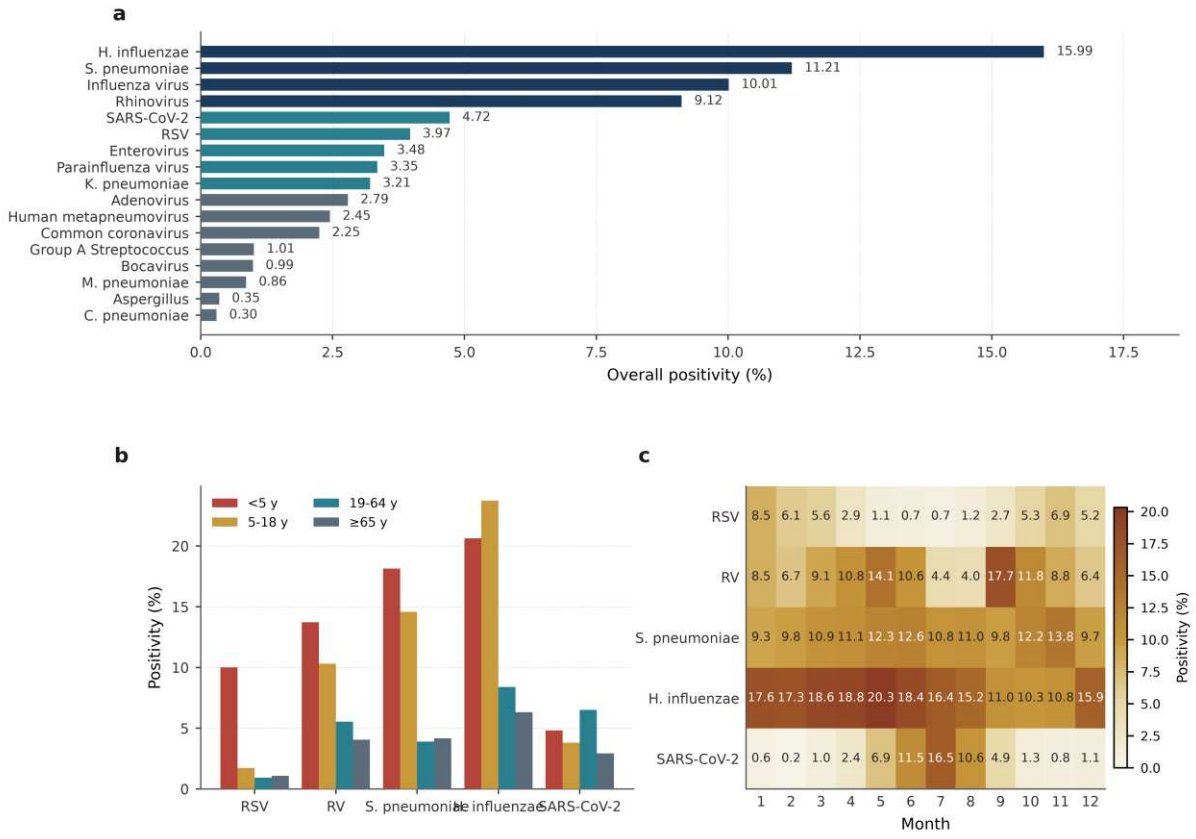



Figure 2 Pathogen distribution across the cohort. (a) Overall positivity of 17 pathogens with positivity ≥ 0.3%; (b) positivity of the five modelled pathogens by age group; (c) monthly distribution of positivity for the five modelled pathogens.

Additional results are provided in the Supplementary Material: city-specific exposure profiles (Table S10, Figure S1), exposure-positivity correlations (Table S11, Figure S2), surveillance-stream stratification (Table S12, Figure S3) and age-stratified associations (Figure S4) (Sections S10–S13). Descriptive weekly time series, contour views of the exposure–lag–response surfaces and model-fit diagnostics are provided as Figures S14 to S16.

### 3.2 Lagged associations from distributed lag non-linear models

Cumulative *RR* estimates for the 90th versus 10th percentile of each exposure are shown in Table 3 and Figure 4, and the corresponding exposure–response and lag-window curves in Figures 5 and 6. The relative-humidity–RSV association is the study's flagship result and, deliberately, its cautionary one. In single-exposure models higher relative humidity was positively associated with RSV positivity (2.011 (*95% CI* 1.060–3.815)) and ozone inversely (0.198 (*95% CI* 0.072–0.544)); the humidity association was the most stable single-exposure finding of the study (to model specification), with direction and magnitude unchanged across all six cross-basis and trend-spline specifications (Table S17) and all six lag parametrisations (Table S6), remaining nominally significant in five of six. It reached nominal significance but did not survive within-pathogen FDR control (*P = 0.032,* adjusted *P = 0.130*), and in the dual-exposure sensitivity analysis it attenuated to the null once temperature was co-adjusted (Section 3.5, Table S8). Thus the association's collapse under co-adjustment, not the association per se, is the substantive finding: it demonstrates how a candidate environmental signal can be fragile when co-varying exposures are ignored. The inverse RSV–O3 association was the only RSV estimate to survive FDR control.

The remaining pathogen–exposure associations in Table 3 are reported as exploratory and should not be read as established effects. Rhinovirus, H. influenzae and S. pneumoniae each showed several nominally significant

associations spanning temperature, humidity and combustion-related pollutants (e.g. H. influenzae with temperature 2.452 (*95% CI* 1.162–5.175), humidity 0.493 (*95% CI* 0.360–0.676), PM2.5 0.500 (*95% CI* 0.330–0.757) and PM10 0.600 (*95% CI* 0.417–0.864); S. pneumoniae with temperature 0.418 (*95% CI* 0.176–0.989) and CO 2.692 (*95% CI* 1.461–4.961); SARS-CoV-2 with CO 6.582 (*95% CI* 2.198–19.707) and NO2 3.246 (*95% CI* 1.263–8.344)), but every temperature estimate was unstable across trend-spline specifications (Section 3.5, Table S17) and is flagged exploratory wherever it appears. These patterns are presented to illustrate the screening output, not as confirmed pathogen-specific drivers.

| Pathogen | Temperature | Relative humidity | PM2.5 | PM10 | SO2 | NO2 | CO | O3 |
|---|---|---|---|---|---|---|---|---|
| RSV | 0.49 | 2.01* | 0.75 | 0.73 | 0.33 | 0.52 | 0.98 | 0.20† |
| Rhinovirus | 3.57† | 1.77† | 1.26 | 1.25 | 0.38† | 2.88† | 2.52† | 0.69 |
| S. pneumoniae | 0.42* | 1.36 | 1.37 | 1.41 | 1.48 | 0.69 | 2.69† | 0.72 |
| H. influenzae | 2.45† | 0.49† | 0.50† | 0.60† | 0.81 | 1.15 | 0.65 | 1.63* |
| SARS-CoV-2 | 1.65 | 1.06 | 0.45 | 0.55 | 0.93 | 3.25* | 6.58† | 0.56 |

Table 3 Cumulative relative risk (*RR*) for the 90th versus 10th percentile of each exposure over the lag 0–6 week window, by pathogen. All five pathogens retained for DLNM are shown. Symbols: * *95% CI* excludes 1 (nominal 5% level); † significant after Benjamini–Hochberg false discovery rate control within pathogen (12 of 40). BH-adjusted *P* values of the surviving estimates: RSV-O3 0.014; rhinovirus-temperature/relative humidity/SO2/NO2/CO 0.013/0.013/0.013/0.012/0.013; S. pneumoniae-CO 0.012; H. influenzae-relative humidity/PM2.5/PM10/temperature <0.001/0.004/0.016/0.037; SARS-CoV-2-CO 0.006. The humidity-RSV estimate (nominal *P = 0.032*) did not survive the correction (adjusted *P = 0.130*). †† Temperature estimates were unstable across trend-spline specifications (cumulative *RR* range 0.172-0.505) and are exploratory only; see Section 3.5 and Table S17.

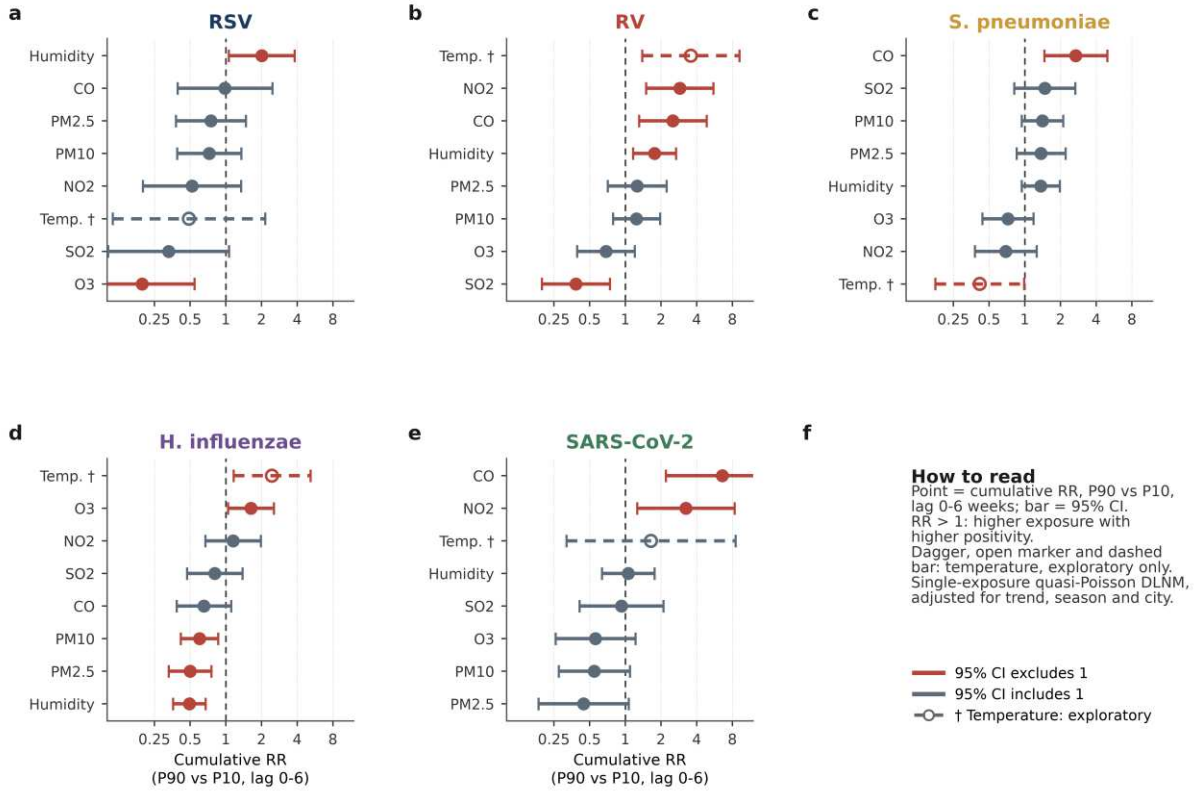



Figure 4 Cumulative relative risks (*95% CI*) for the 90th versus 10th percentile of eight exposures over lag 0–6 weeks, by pathogen. Red markers indicate estimates whose *95% CI* excludes 1. Temperature estimates (T) are drawn with open markers and dashed intervals and are exploratory only: they were unstable across trend-spline specifications (cumulative *RR* range 0.172–0.505 for RSV; Section 3.5, Table S17).

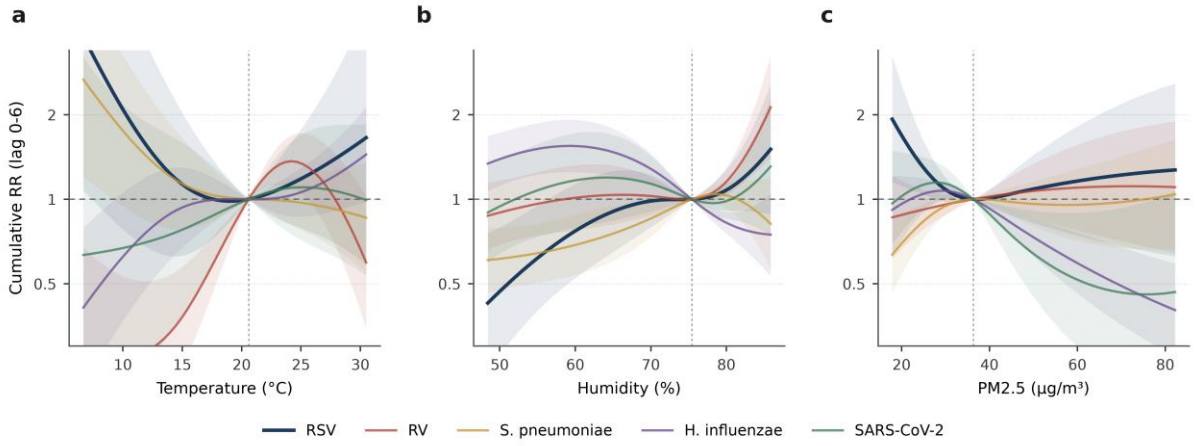



Figure 5 Cumulative exposure–response curves for the five modelled pathogens over weekly mean temperature (a), relative humidity (b) and PM2.5 (c), with 95% confidence bands over the lag 0–6 week window and the median reference value indicated by the dotted line. Temperature panels are shown for completeness; these estimates were sensitive to trend-spline degrees of freedom and should be read as exploratory (Section 3.5, Table S17).

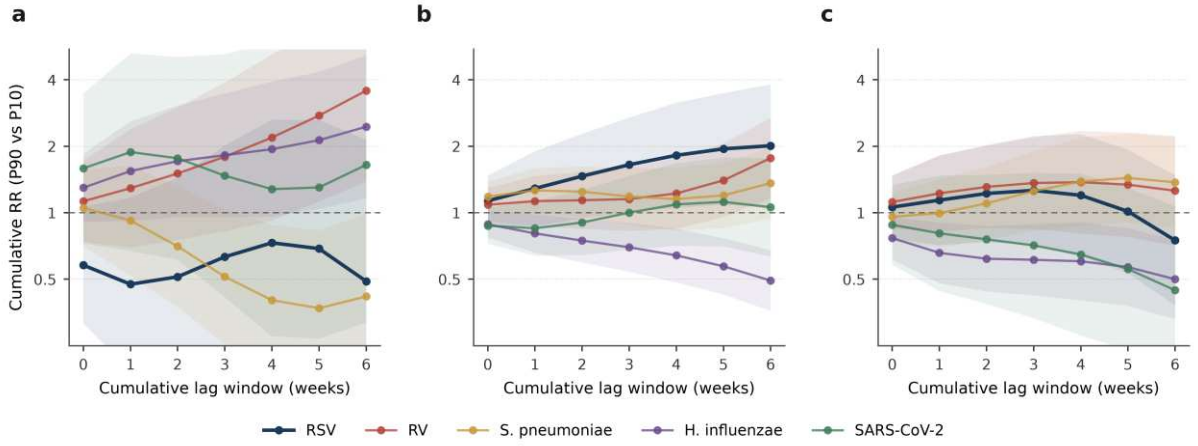



Figure 6 Cumulative relative risk (*95% CI*) for the 90th versus 10th percentile of temperature, relative humidity and PM2.5 over increasing lag windows (lag 0–1 through lag 0–6 weeks), by pathogen. Temperature estimates are shown for completeness and are exploratory (Section 3.5, Table S17).

### 3.3 Joint exposure–lag–response structure and model fit

The cumulative exposure–lag–response surfaces for the four primary pathogen–exposure pairs (Figure 9; contour form in Figure S15) show that the associations reported in Table 3 accumulated over the lag window rather than arising from a single week. When the exposure was sustained for one week the *RR* stayed close to the null in every panel, and it departed from 1 progressively as more lag weeks shared the same exposure: for RSV and relative humidity the exposure-axis range widened from 0.96-1.16 at lag week 0 to 0.43-1.51 at lag week 6, and for H. influenzae and temperature from 0.93-1.19 to 0.41-1.44. The full 0-6 week window therefore reproduces the cumulative estimates of Table 3, and the isolated lag strata do not. The shape of the build-up differed between pairs: RSV and both H. influenzae signals widened monotonically with lag, whereas the cold-preferring S. pneumoniae temperature response peaked at lag week 5 (*RR* 3.00) and eased slightly at week 6, indicating that its effect saturates within the window. The direction of the gradient was uniform across the window for every pair. This pattern follows from the cross-basis parametrisation (exposure *df* 3 x lag *df* 3), in which single-week contrasts are weakly identified (Section 2.4), and the surfaces thus provide a visual check that the reported cumulative estimates are not driven by any single lag stratum.

Figure S16 plots observed against fitted positivity for all five modelled pathogens, each against its strongest exposure. Agreement between the two is high in every case ($r$ = *0.66* to 0.80), indicating that the models reproduce the observed variation adequately. The raw marginal correlation for the same exposures is much weaker (absolute values 0.00 to 0.27) but agrees in direction with the modelled associations.

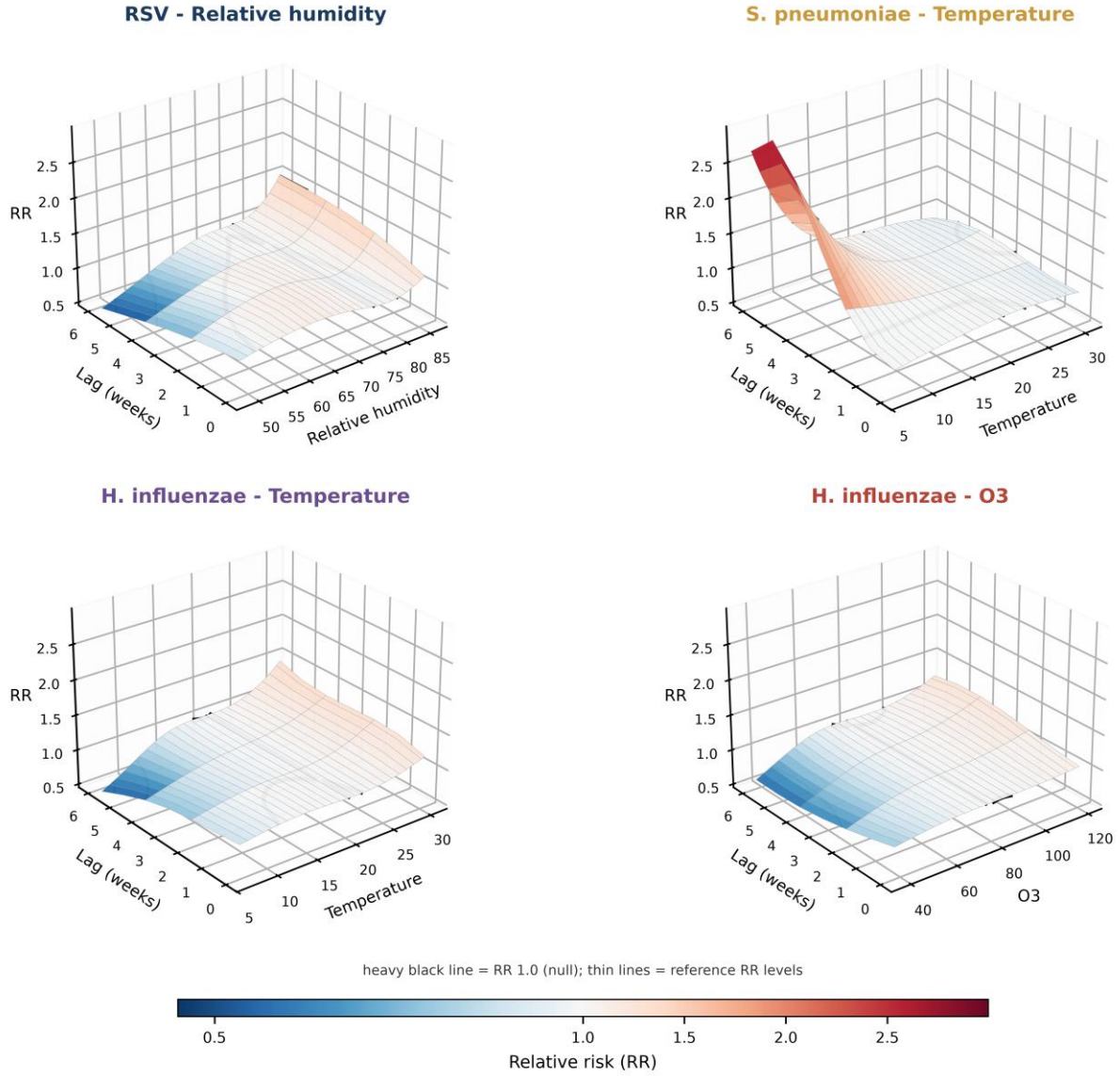



Figure 9 Cumulative exposure–lag–response surfaces for four pathogen–exposure pairs: RSV–relative humidity, S. pneumoniae–temperature, H. influenzae–temperature and H. influenzae–O3. Each surface shows the *RR* for an exposure sustained over the leading 0–L lag weeks (0–6), remaining lags held at the median. The colour scale is shared and centred at *RR = 1*; the heavy black line marks *RR = 1* and thin rings are reference levels, drawn only where the panel reaches them.

Every one of the 40 models used the same 700 city-week observations. This is expected rather than coincidental: all eight exposures are complete over the same weeks, and the only rows excluded are the first six weeks of each city, needed to complete the lag 0–6 window. The rule therefore depends on neither the exposure nor the pathogen. Dispersion, residual autocorrelation and calibration still vary across models, which confirms the 40 fits are distinct (Table S18). The observed/expected ratio is 1.000 throughout, so the fitted series reproduces the observed totals.

### 3.4 Co-detection decomposition

Co-detection was treated throughout as a descriptive attribute rather than a modelled dependency. The decomposition of the pooled estimate into single- and co-detection components (Supplementary Material, Table S9) shows the relative-humidity-RSV association is present, with the same sign, in both detection regimes, and that humidity does not alter the co-detection rate, arguing against a composition artefact; full results are in Section S9.

### 3.5 Sensitivity analysis

Changing the cross-basis degrees of freedom (exposure *df* 3–6, lag *df* 3–4) moved the humidity estimate within a narrow band, indicating that the cross-basis specification was not a major source of variation. Temperature remained the least stable exposure: its cumulative *RR* for RSV ranged from 0.172 to 0.505 across trend degrees of freedom 4–10, with its 95% confidence interval crossing the null relative to the primary specification, which is consistent with collinearity between temperature and seasonality. Humidity estimates kept their direction and magnitude across all specifications (*RR* 1.444–2.477); they remained nominally significant in five of the six specifications, losing significance only under the most flexible specification (6-4-10), and are reported as a comparatively stable single-exposure signal (Table S17).

Two further sensitivity analyses are reported in the Supporting Information: (i) autocorrelation-robust (cluster-by-city and Newey–West HAC) standard errors, computed for all 12 pathogen–exposure estimates that survived within-pathogen FDR control using a reduced-form quasi-Poisson panel model (linear exposure, city fixed effects, natural cubic spline trend *df = 8*). Six of the 12 retained a *95% CI* excluding 1 under both cluster-by-city and Newey–West estimators, confirming that residual autocorrelation does not manufacture the FDR-surviving signals, although it narrows the subset that holds under the reduced-form linear specification; and (ii) variance inflation factors showing substantial collinearity among combustion-related pollutants, supporting the single-exposure design.

To probe the robustness boundary of the single-exposure estimates, a dual-exposure DLNM was fitted for selected pathogen-exposure pairs (RSV humidity with temperature or PM2.5; S. pneumoniae and H. influenzae temperature with a second exposure), entering two cross-basis terms together with the city and trend terms on the 700 city-week observations. This analysis marks the boundary of what the data can support: once humidity and temperature were entered together, the RSV–humidity association attenuated from *RR* 2.01 to 1.05 and no longer excluded the null, and the temperature term moved from *RR* 0.49 to 1.03; the same attenuation occurred for every pathogen-exposure pair examined, and none remained significant after mutual adjustment. With the two cross-basis terms contributing 18 exposure parameters, plus city and trend terms, on 700 observations the dual-exposure models are over-parameterised, so these point estimates should not be over-interpreted in either direction. They are reported as a sensitivity analysis (Table S8) and should be read as demonstrating that the single-exposure estimates are not separable from the co-varying exposures in this panel: the principal reason we describe the single-exposure results as candidate indicators rather than established effects.

## 4 Discussion

In this multi-pathogen sentinel analysis of 23 014 ARI cases across seven hospitals in Sichuan Province, meteorological conditions and air pollutants were associated with pathogen positivity in a lagged, cumulative and pathogen-specific manner. We first summarise the principal results, then set out the transferable contribution. The specific associations reported here are, as Section 3.5 shows, largely not separable from co-varying exposures in a panel of this size; the incremental value therefore lies less in any single pathogen-exposure pair than in the screening sequence itself: pre-specified eligibility criteria, single-exposure distributed lag models as the first pass, within-pathogen false discovery rate control, systematic specification sensitivity, dual-exposure models marking the boundary of what the data can support, and decomposition of the co-

detection mixture. Each step answers a specific inferential threat (selection, multiplicity, specification, collinearity and composition) and each is reproducible from routine sentinel data. We offer this sequence as a template for other multi-pathogen sentinel networks. Three elements distinguish this work from the single-pathogen distributed-lag studies that dominate the field. First, the screening sequence is applied across 40 exposure–pathogen pairs under within-pathogen multiplicity control, so candidate signals are ranked rather than reported in isolation. Second, exposure collinearity is treated as a primary inferential threat: the humidity–RSV signal survives specification sensitivity yet collapses under temperature co-adjustment, a fragility that single-exposure analyses would have reported as real. Third, opposing species-level bacterial and viral environmental gradients are resolved within one framework rather than at the syndrome level. Together, these are what set the study apart from the dominant single-pathogen design.

Environmental sensitivities diverged sharply across the five pathogens rather than moving as a single respiratory class. RSV tracked high relative humidity, an effect accumulating over the six-week lag window, whereas its temperature association was non-significant (*RR* 0.489, *95% CI* 0.112-2.147) and S. pneumoniae showed a cold preference that likewise did not survive within-pathogen FDR control; both are exploratory. Rhinovirus responded to the broadest set of exposures (high temperature, humidity, NO2 and CO; low SO2), and SARS-CoV-2 tracked combustion pollutants (CO, NO2) rather than meteorology. H. influenzae ran opposite to the viruses on the axes that survived multiplicity control: more prevalent at higher temperature and less at higher humidity, PM2.5 and PM10. This bacterial-versus-viral contrast matches a multicentre Chinese pneumonia study in which bacterial detection rose with temperature and fell with pollutants while viral detection showed the reverse[18], and is biologically plausible because enveloped viruses are stabilised by cold, dry air whereas many bacteria tolerate warmth[5, 6]. The same humidity-RSV pattern has been reported in Suzhou and Singapore[4, 35]; the divergence argues against pooling bacteria and viruses into one "respiratory pathogen" exposure class and against seasonality alone as a basis for early warning.

The lag structure is central to interpretation. Effects for humidity were weak at short lags and only became statistically distinguishable as the lag window extended to six weeks. Analyses restricted to concurrent-week contrasts would therefore have underestimated the association. This reinforces the argument for DLNM-type frameworks in pathogen surveillance, where the relevant biological timescale is measured in weeks rather than days[36].

Young children carried the greatest burden: under-5s had an RSV positivity nearly eleven times that of adults aged 19-64 years, and their humidity response exceeded that of older groups[13, 14], consistent with the established paediatric epidemiology of RSV[37]. From a surveillance perspective, environmental indicators should be weighted by age composition and paediatric capacity anticipated during humid, cold periods. The pollutant signals are directionally consistent with the wider pneumonia literature[38, 39, 40], but the single-exposure design cannot isolate any one pollutant, so direction alone argues against a purely artefactual result rather than establishing a causal pathway.

Two pathogens of clinical importance were excluded for reasons that lie in the surveillance design rather than the pathogens: influenza virus, despite its high case count, was strongly over-dispersed (dispersion 5.12), and enterovirus contributed only 758 positive cases in the analytic panel, below the pre-specified minimum. Both remain visible in Table 1 and Figure 2, so that the omission is explicit rather than silent.

The decomposition also marks the boundary of a single-outcome model: neither stratum excludes confounding by whatever drives co-detection, and the disagreement between the two S. pneumoniae–temperature strata reinforces the exploratory status already assigned to temperature. Settling these questions would require a joint multi-pathogen model with specimen-level exposure histories, which we regard as the most important methodological extension of this work.

Several limitations warrant emphasis, and they shape how the findings should be interpreted.

First, the outcome is pathogen positivity among specimens tested at sentinel hospitals, not population incidence. Cases were ascertained among patients who presented to and were tested at these facilities, so both care-seeking behaviour and testing practice influence the observed positivity. Positivity is conditional on presentation and on the testing decision, and the estimates must not be read as population-level infection rates or used to infer community burden. The same consideration applies to the stream comparison: ILI and SARI capture different case definitions, severities and age distributions, and the higher SARI positivity for RSV partly reflects the concentration of young children among hospitalised cases.

Second, exposure data were gridded reanalysis products (ERA5 and CAMS) without calibration to local monitoring stations. As detailed in Section 2.1, two sources of exposure error (grid-placement error and the systematic reanalysis-versus-station difference) were separated; the latter cannot be quantified without station records, so pollutant estimates carry an additional, only partially quantifiable uncertainty beyond sampling error, which is why we treat them as exploratory. Modelled concentrations also differ in magnitude from station measurements, so we emphasise the direction of associations and the shape of the lag-accumulation pattern rather than absolute *RR* values; pollutant relative risks are directional indicators only, not dose-response parameters[28].

Third, all models were single-exposure, and the exposures are strongly inter-correlated and seasonally co-varying, so the independent contribution of any single pollutant could not be isolated and a single-exposure estimate may proxy for another exposure or for season. This is a statistical rather than substantive limitation: the variables act jointly, but the present sample and collinearity do not permit their separate estimation with confidence. The inverse NO2-RSV association illustrates the point (counter-intuitive, not FDR-robust, reported as exploratory); temperature associations were likewise unstable across trend-spline specifications (*RR* 0.172-0.505), whereas humidity held its direction and magnitude across all six (*RR* 1.444-2.477) and is therefore designated the most stable single-exposure signal.

Fourth, the ≥65-year subgroup contributed few positive cases and its estimates were imprecise; the age-stratified results for this group are exploratory only and do not support definitive inference. Fifth, the ecological design aggregated individual cases to the city-week level, precluding individual-level inference and leaving open confounding by unmeasured factors such as vaccination coverage, school calendars and health-seeking behaviour. Sixth, lag-1 autocorrelation of Pearson residuals ranged from 0.16 to 0.36 across models, indicating residual temporal structure that the trend and seasonal splines did not fully absorb; confidence intervals may therefore be somewhat narrow. Seventh, the two-year study window captures only two seasonal cycles per city, which precludes assessment of inter-annual variability and limits the stability of the season-adjusted estimates; extension to a multi-year series is required before any operational use. Finally, reanalysis data represent modelled rather than measured exposure, which should be borne in mind when comparing effect magnitudes with station-based studies.

## 5 Conclusions

In this sentinel-based analysis of 23 014 ARI cases, meteorological conditions and air pollutants were associated with pathogen positivity in a lagged, cumulative and pathogen-specific manner. The most stable single-exposure signal (to model specification) was a positive, cumulative association between relative humidity and RSV positivity, which was strongest among children under five years and in the inpatient SARI stream but attenuated to the null once temperature was co-adjusted; it therefore requires prospective confirmation in independent sentinel networks before any early-warning application. For H. influenzae the humidity and particulate-matter associations survived multiplicity control, and for S. pneumoniae the carbon-

monoxide association did so; by contrast, the temperature associations (including the cold-preferring S. pneumoniae estimate) did not survive FDR control. They were further sensitive to model specification and, in dual-exposure sensitivity analyses, reversed direction depending on the co-adjusted exposure, and are reported only as exploratory findings that should not be interpreted as established effects. Given the test-conditional nature of the outcome (Methods 2.3), the use of uncalibrated reanalysis exposure data and the single-exposure design, these findings identify candidate environmental indicators for multi-pathogen ARI surveillance that warrant prospective validation against station-measured exposure and individual-level data, rather than supporting immediate operational use. Future work should extend the lag window to a full seasonal cycle and to a multi-year series, employ multi-pollutant or distributed-lag interaction models, and test whether environmental signals prospectively improve the timing of alerts. The analytical sequence itself (a priori eligibility criteria, within-pathogen FDR control, specification sensitivity, dual-exposure boundary testing and co-detection decomposition), rather than any single exposure-pathogen pair, is the transferable output of this study.

## CRediT authorship contribution statement

Fengmiao Hu: Conceptualization, Methodology, Formal analysis, Writing – original draft. Xingyu Zhou: Data curation, Formal analysis, Writing – original draft. Lijun Zhou: Data curation, Validation. Zhirui Li: Data curation, Validation. Shuang Dong: Data curation, Validation. Biao Yang: Data curation, Validation. Chongkun Xiao: Conceptualization, Supervision, Writing – review & editing, Funding acquisition.

## Patient and public involvement

This study was based on routine, de-identified sentinel surveillance data. Patients and the public were not involved in the design, conduct, analysis, or reporting of the study.

## Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

The case-level surveillance data are not publicly available because of confidentiality obligations, but the aggregated data and analysis code are available from the corresponding author on reasonable request.

## Access to data

The corresponding author had full access to all the data in the study and takes final responsibility for the decision to submit for publication.

## Funding

This work was supported by the Sichuan Center for Disease Control and Prevention Youth Fund (Grant No. QN202509) and the Chengdu Medical Research Project (Grant No. 2026447). Neither funder had any role in the study design, data collection, analysis, interpretation, or the writing of the report, and neither funder had any role in the decision to submit for publication.

## Acknowledgements

Meteorological data were provided by the ECMWF ERA5 reanalysis through the Copernicus Climate Change Service; air pollutant concentrations were provided by the Copernicus Atmosphere Monitoring Service (CAMS); both were accessed via the Open-Meteo API. We thank the staff of the seven sentinel hospitals and the Sichuan provincial reference laboratory for specimen testing and data curation.

## Ethics approval

The study was approved by the Ethics Review Committee of the Sichuan Center for Disease Control and Prevention (SCCDCIRB No. 2026-026). As a retrospective analysis of de-identified surveillance data, the requirement for individual informed consent was waived by the Ethics Review Committee of the Sichuan Center for Disease Control and Prevention (SCCDCIRB No. 2026-026).

## Prior presentation

This manuscript has not been previously published and is not under consideration elsewhere; the findings have not been presented at a conference.

## References


[1] GBD 2019 Lower Respiratory Infections Collaborators. Age-sex differences in the global burden of lower respiratory infections and risk factors, 1990-2019: results from the Global Burden of Disease Study 2019. Lancet Infect Dis. 2022;22(11):1626-47. doi:10.1016/S1473-3099(22)00510-2.

[2] Metersky ML, Masterton RG, Lode H, File TM, Babinchak T. Epidemiology, microbiology, and treatment considerations for bacterial pneumonia complicating influenza. Int J Infect Dis. 2012;16(5):e321-31. doi:10.1016/j.ijid.2012.01.003.

[3] Woodhead M, Blasi F, Ewig S, et al. Guidelines for the management of adult lower respiratory tract infections. Eur Respir J. 2005;26(6):1138-80. doi:10.1183/09031936.05.00055705.

[4] Lu Y, Ren S, Shao X, Tian J, Hu F, Yao F, Zhang T, Zhao G. Association of ambient temperature and relative humidity with respiratory syncytial virus infections among hospitalized children in Suzhou, eastern China: a time-series analysis. GeoHealth. 2025;9(5):e2025GH001353. doi:10.1029/2025GH001353.

[5] Paynter S. Humidity and respiratory virus transmission in tropical and temperate settings. Epidemiol Infect. 2014;143(6):1110-8. doi:10.1017/S0950268814002702.

[6] Shaman J, Kohn M. Absolute humidity modulates influenza survival, transmission and seasonality. Proc Natl Acad Sci U S A. 2009;106(9):3243-8. doi:10.1073/pnas.0806852106.

[7] Yang J, Yang Z, Li Q, Li M, Liu D, Liu X, Tong S, Sun Q, Feng L, Ou CQ, Liu Q. Influence of air pollution on influenza-like illness in China: a nationwide time-series analysis. EBioMedicine. 2023;87:104421. doi:10.1016/j.ebiom.2022.104421.

[8] Pan Y, Yao L, Huang B, et al. Time series analysis of the impact of air pollutants on influenza-like illness in Changchun, China. BMC Public Health. 2025;25:1456. doi:10.1186/s12889-025-22110-x.

[9] Romieu I, Samet J, Smith K. Outdoor air pollution and acute respiratory infections among children in developing countries. J Occup Environ Med. 2002;44(7):640-9. doi:10.1097/00043764-200207000-00010.

[10] Gasparrini A, Armstrong B, Kenward MG. Distributed lag non-linear models. Stat Med. 2010;29(21):2224-34. doi:10.1002/sim.3940.

[11] Gasparrini A, Leone M. Attributable risk from distributed lag non-linear models. BMC Med Res Methodol. 2014;14:55. doi:10.1186/1471-2288-14-55.

[12] Armstrong B. Time series methods. In: Elliott P, Wartenberg D, eds. Spatial Epidemiology. Oxford: Oxford University Press; 2004.

[13] Nair H, Nokes DJ, Gessner BD, et al. Global burden of acute lower respiratory infections due to respiratory syncytial virus in young children: a systematic review and meta-analysis. Lancet. 2010;375(9725):1545-55. doi:10.1016/S0140-6736(10)60206-1.

[14] Shi T, McAllister DA, O'Brien KL, et al. Global, regional, and national disease burden estimates of acute lower respiratory infections due to respiratory syncytial virus in young children in 2015: a systematic review and modelling study. Lancet. 2017;390(10098):946-58. doi:10.1016/S0140-6736(17)30938-8.

[15] Jain S, Self WH, Wunderink RG, et al. Community-acquired pneumonia requiring hospitalization among U.S. adults. N Engl J Med. 2015;373(5):415-27. doi:10.1056/NEJMoa1500245.

[16] Liu X, Kang M, Ma WJ, et al. Effect of meteorological factors on risk of influenza-like illness in Guangzhou: a time-series analysis. J Environ Hyg (Chin). 2018;8(5):374-80. doi:10.13421/j.cnki.hjwsxzz.2018.05.003.

[17] Wang YY, Xu L, Hang RF, et al. Time series analysis of the association between temperature and incidence of influenza-like illness among residents in Huai'an city, Jiangsu province. Chin J Public Health. 2025;41(9):1031-5. doi:10.11847/zgggws1146908.

[18] Liu P, Li H, Liu Y, et al. Incidence, etiology, and environmental risk factors of community-acquired pneumonia requiring hospitalization in China: a 3-year, prospective, age-stratified, multicenter case-control study. Open Forum Infect Dis. 2021;8(11):ofab499. doi:10.1093/ofid/ofab499.

[19] Wang ZB, Ren L, Lu QB, Zhang XA, Miao D, Hu YY, Dai K, Li H, Luo ZX, Fang LQ, Liu EM, Liu W. The impact of weather and air pollution on viral infection and disease outcome among pediatric pneumonia patients in Chongqing, China, from 2009 to 2018: a prospective observational study. Clin Infect Dis. 2021;73(2):e513-e522. doi:10.1093/cid/ciaa997.

[20] Hu F, Zhou X, Zhou L, Li Z, Dong S, Xiao C. Analysis of multiplex pathogen sentinel surveillance results of SARS-CoV-2 and influenza virus in Sichuan Province, 2024-2026. Front Public Health. 2026;14:1961515. doi:10.3389/fpubh.2026.1961515.

[21] Hu F, Zhou X, Zhou L, Li Z, Dong S, Xiao C. Pathogen prevalence, feature composition and cross-centre generalisability of machine learning diagnostic models for multi-pathogen respiratory infection. Front Public Health. 2026;14:1961816. doi:10.3389/fpubh.2026.1961816.

[22] Hu F, Zhou X, Zhou L, Li Z, Dong S, Xiao C. The sentinel surveillance results of human respiratory syncytial virus in Sichuan Province, 2024-2025. J Trop Dis Parasitol. 2026;24(4):225-9. doi:10.20199/j.issn.1672-2302.2026.04.007.

[23] von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. PLoS Med. 2007;4(10):e296. doi:10.1371/journal.pmed.0040296.

[24] Chinese Center for Disease Control and Prevention. Notice on printing and issuing the Technical Protocol for National Sentinel Surveillance of Acute Respiratory Infectious Diseases (trial version), No. A3361. Beijing: Chinese Center for Disease Control and Prevention; 2024 Dec 4. [in Chinese]

[25] Hersbach H, Bell B, Berrisford P, et al. The ERA5 global reanalysis. Q J R Meteorol Soc. 2020;146(730):1999-2049. doi:10.1002/qj.3803.

[26] Inness A, Ades M, Agustí-Panareda A, et al. The CAMS reanalysis of atmospheric composition. Atmos Chem Phys. 2019;19(6):3515-56. doi:10.5194/acp-19-3515-2019.

[27] Copernicus Atmosphere Monitoring Service. CAMS global atmospheric composition forecasts: near-real-time data documentation. ECMWF. https://ads.atmosphere.copernicus.eu/datasets/cams-global-atmospheric-composition-forecasts.

[28] Dufek A, Ambrizzi T, Rocha R, et al. Are reanalysis data useful for calculating climate indices over South America? Ann N Y Acad Sci. 2008;1146(1):87-104. doi:10.1196/annals.1446.010.

[29] Ding S, Liu D. Evaluation of the CAMS reanalysis for atmospheric black carbon and carbon monoxide over the north China plain. Environ Pollut. 2022;314:120286. doi:10.1016/j.envpol.2022.120286.

[30] Li W, Wang Y, Yi Z, Guo B, Chen W, Che H. Evaluation of MERRA-2 and CAMS reanalysis for black carbon aerosol in China. Environ Pollut. 2024;343:123182. doi:10.1016/j.envpol.2023.123182.

[31] Ma L, Wei Y, Dong W. Evaluation and spatiotemporal evolution of surface air temperature estimated by ERA5-Land reanalysis data in East China. J Appl Meteorol Climatol. 2025;64(8):943-58. doi:10.1175/JAMC-D-24-0108.1.

[32] Hayden RT, Gu Z, Rodriguez A, Tanioka L, Ying C, Morgenstern M, Bankowski MJ. Comparison of two broadly multiplexed PCR systems for viral detection in clinical respiratory tract specimens from immunocompromised children. J Clin Virol. 2012;53(4):308-13. doi:10.1016/j.jcv.2011.12.020.

[33] Gasparrini A. dlnm: distributed lag non-linear models. R package version 2.4.10. 2023. https://CRAN.R-project.org/package=dlnm.

[34] Gasparrini A, Armstrong B, Kenward MG, Michelozzi P. Reducing and meta-analysing estimates from distributed lag non-linear models. BMC Med Res Methodol. 2013;13:1. doi:10.1186/1471-2288-13-1.

[35] Lee MH, Mailepessov D, Yahya K, Loo LH, Maiwald M, Aik J. Air quality, meteorological variability and pediatric respiratory syncytial virus infections in Singapore. Sci Rep. 2023;13:1001. doi:10.1038/s41598-022-26184-0.

[36] Tamerius J, Nelson M, Zhou S, et al. Global influenza seasonality: reconciling patterns across temperate and tropical regions. Environ Health Perspect. 2011;119(4):439-45. doi:10.1289/ehp.1002383.

[37] Buchan SA, Chung H, To T, Daneman N, Guttmann A, Kwong JC, et al. Estimating the incidence of first respiratory syncytial virus hospitalization in children born in Ontario, Canada. J Pediatric Infect Dis Soc. 2023;12(7):421-30. doi:10.1093/jpids/piad045.

[38] Zheng P, Wang J, Zhang Z, et al. Air pollution and hospital visits for acute upper and lower respiratory infections among children in Ningbo, China: a time-series analysis. Environ Sci Pollut Res. 2017;24(23):18860-9. doi:10.1007/s11356-017-9279-8.

[39] Zhang W, Ling J, Zhang R, et al. Short-term effects of air pollution on hospitalization for acute lower respiratory infections in children: a time-series analysis study from Lanzhou, China. BMC Public Health. 2023;23:1145. doi:10.1186/s12889-023-16533-7.

[40] Shi W, Liu C, Annesi-Maesano I, et al. Ambient PM2.5 and its chemical constituents on lifetime-ever pneumonia in Chinese children: a multi-center study. Environ Int. 2021;146:106176. doi:10.1016/j.envint.2020.106176.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupportingInformationSichuanARIDLNM.docx