Journal of Information Organization

DLINE JOURNALS

Print ISSN : 2278 – 6503
Online ISSN: 2278 – 6511

JIO 2026: 16 (4)
https://doi.org/10.6025/jio/2026/16/4/203-236

Mapping the Global Landscape of Respiratory Virus Co-circulation: Temporal Dynamics, Pairwise Associations, and Multi-Pathogen Detection Patterns

Yue Yong
Department of Computer Science and Technology
University of Bedfordshire
United Kingdom

ABSTRACT

Background: *Respiratory viruses circulate as an interconnected multi-pathogen system, yet most surveillance analyses characterise individual pathogens in isolation. Integrated, missingness-aware descriptions of country-level co-detection remain limited, and ecological associations are frequently misinterpreted as biological interaction.*

Methods: *A retrospective, observational, ecological analysis was conducted on international weekly respiratory-virus surveillance from 184 countries (4 January 2021 – 12 October 2026). Of 72,598 raw records, 55,742 contained positive specimen counts (64,677,119 specimens); consolidation of 23,739 repeated country-week keys yielded 41,465 analytical country-week units. Ten virus groups were examined. Missing pathogen counts were retained as missing (mean missingness 61.8%) rather than recoded as zero. Analyses comprised pathogen-specific detection proportions with virus-specific denominators; pairwise 2×2 contingency tables, odds ratios (OR) and Fisher's exact tests with Benjamini–Hochberg false-discovery-rate control across 45 pairs; Spearman correlations of log1p-transformed counts; monthly detections per 1,000 specimens; and multi-pathogen burden restricted to country-weeks with e"5 reported virus groups. Weighted network construction and confounding-adjusted modelling are specified as prospective extensions.*

Results: *Influenza A showed the highest positive proportion (78.49%; 25,986/33,108) and cumulative detections (4,395,375), followed by RSV (63.51%; 17,678/27,834); bocavirus was lowest (24.59%; 2,297/ 9,341). Rhinovirus had the highest median positive count (13). The strongest co-detection association was adenovirus–metapneumovirus (OR = 20.33; 4,387 co-positive of 11,292 complete pairs). All rank correlations were positive, strongest for adenovirus–parainfluenza (ñ = 0.777), adenovirus–metapneumovirus (ñ = 0.753) and metapneumovirus–parainfluenza (ñ = 0.739), and weakest for bocavirus–influenza B (ñ = 0.240). Among 12,441 eligible country-weeks, 9,932 (79.83%) reported e"2 virus groups and 339 (2.72%) reported all ten.*

Conclusions: *Multi-pathogen co-detection is a prominent feature of global surveillance, with a consistently correlated adenovirus–metapneumovirus–parainfluenza–rhinovirus grouping. Because denominators are unequal, missingness is substantial and estimates are unadjusted, these findings represent country-week ecological associations—not comparable prevalence estimates, individual-level coinfection, or evidence of*

viral interference. The framework provides a defensible basis for adjusted, temporal and network-based investigation.

Keywords: Respiratory Viruses, Co-Circulation, Co-Detection, Global Surveillance, Ecological Analysis, Odds Ratio, Spearman Correlation, Multi-Pathogen Burden, Pathogen Association Network, Missing Data, Influenza A, Respiratory Syncytial Virus

Received: 4 March 2026, Revised: 6 June 2026, Accepted: 25 June 2026

Copyright: DLINE

1. Introduction

Seasonal respiratory infections continue to represent a substantial challenge to global public health, with their epidemiological dynamics influenced by complex interactions among viral evolution, population immunity, environmental conditions, and temporal variations in pathogen circulation. Seasonal influenza, in particular, remains a major contributor to the global burden of respiratory disease, accounting for approximately 500,000 deaths annually worldwide [1]. Its transmission dynamics are shaped by the continuous antigenic evolution of influenza A (IAV) and influenza B (IBV) viruses, whose antigenic drift facilitates immune evasion and contributes to the complexity of seasonal epidemics [2] (Petrova VN.). More broadly, respiratory pathogens exhibit distinct seasonal characteristics. Respiratory syncytial virus (RSV), for example, predominantly circulates during winter in temperate regions and is an important cause of lower respiratory tract infections, particularly among children younger than two years [3]. In contrast, *Mycoplasma pneumoniae* (MP) has been associated primarily with springtime seasonal peaks ([4]. These differences highlight the heterogeneous seasonal behaviour of respiratory pathogens and the importance of investigating their circulation within a broader epidemiological framework.

Understanding respiratory-virus seasonality requires analytical approaches capable of capturing both regular periodic patterns and complex nonlinear temporal variations. Previous studies have frequently employed wavelet transforms to characterize time-dependent fluctuations in infectious disease activity [5]. Other investigations have incorporated adaptive Fourier decomposition (AFD), a signal-processing technique designed to extract meaningful frequency components from complex signals [6]. In particular, a divide-and-conquer analytical strategy has been proposed to improve the characterization of respiratory disease dynamics. This approach initially applies seasonal-trend decomposition using LOESS (STL) to separate long-term trends from relatively stable seasonal components [7, 8]. Subsequently, AFD is employed to extract meaningful transient components from high-frequency residual signals, thereby improving the representation of complex temporal variations. Such combined approaches offer the potential to characterize nonlinear epidemic behaviour more effectively and have demonstrated accuracy and computational efficiency in influenza antigen variation prediction [9]. These methodological developments emphasize the importance of integrating complementary temporal analysis techniques when investigating the changing dynamics of respiratory pathogens.

2. Background

Despite advances in respiratory virus surveillance, much of the post-pandemic literature has concentrated on changes in individual virus detection rates, epidemic timing, and age-specific distributions [10]. Although these investigations provide important insights into changes in pathogen circulation, comparatively limited attention has been directed towards understanding how respiratory virus co-detection patterns and dominant pairwise associations have been reorganized across the pre-pandemic, pandemic, and post-pandemic periods. This represents an important epidemiological research gap because respiratory virus co-circulation patterns may reflect temporal overlap between pathogens, changes in the populations undergoing testing, and variations in the viral communities identified through multiplex diagnostic assays [11]. Consequently, examining the

relationships among multiple respiratory pathogens, rather than investigating each virus independently, can provide a broader understanding of changes in respiratory-virus surveillance patterns.

The circulation of rhinovirus provides an additional perspective for understanding these temporal changes. Previous investigations have suggested that rhinovirus maintained relatively strong circulation during periods of non-pharmaceutical interventions and frequently re-emerged earlier than several enveloped respiratory viruses [12, 13, 14, 15]. These observations suggest that the persistence of rhinovirus may have contributed to repeated temporal overlap with other respiratory pathogens during changing epidemiological conditions. Furthermore, viral interference has been proposed as a potential mechanism influencing the circulation and interactions of respiratory viruses [16, 17] However, surveillance-based co-detection analyses cannot independently establish such biological mechanisms. Therefore, observed associations between respiratory viruses should be interpreted primarily as statistical and descriptive epidemiological patterns rather than direct evidence of biological interaction. [18, 19].

Against this background, investigating the temporal dynamics of respiratory-virus co-circulation and the structure of pairwise associations is important for understanding the changing landscape of respiratory disease surveillance. A systematic assessment of multi-pathogen detection patterns can help identify frequently co-detected virus pairs, characterize variations in association strength, and reveal changes in the composition of respiratory-virus networks across different epidemiological periods. Such an approach provides a complementary perspective to conventional single pathogen surveillance analyses and can contribute to a more integrated understanding of global respiratory virus circulation.

2.1 Research Problem

Respiratory viruses represent a major and persistent public-health challenge because multiple viral pathogens circulate simultaneously, exhibit heterogeneous seasonal patterns, and contribute to respiratory disease activity across different geographical regions. Although individual respiratory viruses have been extensively investigated, understanding their circulation as an interconnected multi-pathogen system remains challenging. Changes in the detection of one pathogen may coincide with changes in other pathogens because of shared environmental conditions, seasonal influences, geographical distributions, population susceptibility, and variations in surveillance intensity. Consequently, examining each pathogen independently may provide an incomplete representation of the broader respiratory-virus circulation landscape.

The increasing availability of international respiratory-virus surveillance data provides opportunities to investigate these relationships through integrated statistical and temporal analyses. However, surveillance datasets are frequently characterized by heterogeneous reporting practices, unequal pathogen-specific testing coverage, missing observations, and differences in the geographical and temporal representation of participating countries. These characteristics complicate the interpretation of pathogen detection frequencies and pairwise associations. In particular, simultaneous detection of two viruses within the same country-week may indicate overlapping circulation, but it cannot establish that the viruses infected the same individuals or interacted biologically.

The analysis considered in this study uses weekly respiratory-virus surveillance observations from 184 countries between January 2021 and October 2026. It examines ten respiratory-virus groups: influenza A, influenza B, adenovirus, bocavirus, human coronavirus, metapneumovirus, parainfluenza, rhinovirus, respiratory syncytial virus (RSV), and other respiratory viruses. The analytical dataset contains 41,465 consolidated country-week units following the exclusion of records with zero or missing specimen counts and the aggregation of repeated country-week records. The analysis also identifies substantial pathogen-level missingness, with mean missingness of approximately 61.8%, highlighting the need for careful treatment of unreported pathogen observations.

Preliminary findings demonstrate heterogeneous pathogen-specific detection frequencies, strong pairwise associations, and considerable multi-pathogen detection at the country-week level. For example, adenovirus and metapneumovirus exhibit a reported co-detection odds ratio of 20.33, while adenovirus and parainfluenza

demonstrate a Spearman correlation coefficient of 0.777. Furthermore, the co-circulation analysis indicates that 79.83% of eligible country-week observations contain at least two detected virus groups. These findings demonstrate the value of examining respiratory-virus circulation through an integrated multi-pathogen perspective.

Nevertheless, the existing results are primarily descriptive and associational. The reported pairwise odds ratios and correlations are not adjusted for all potentially influential factors, including country specific characteristics, seasonality, temporal trends, surveillance intensity, and differences in reporting completeness. Therefore, the observed associations may partly reflect shared epidemiological or surveillance conditions rather than pathogen-specific relationships.

The central research problem is therefore the absence of a sufficiently integrated, missingness-aware, and surveillance-adjusted analytical framework for characterizing global respiratory-virus co-circulation and distinguishing robust pathogen association patterns from relationships potentially attributable to temporal, geographical, and surveillance-related confounding.

This study addresses the problem by combining pathogen-specific descriptive analysis, pairwise co-detection statistics, correlation analysis, temporal profiling, multi-pathogen burden assessment, and a proposed extension toward adjusted association modelling and network-based characterization. The objective is to develop a systematic understanding of the structure and temporal organization of respiratory-virus co-circulation while maintaining an appropriate distinction between ecological statistical association and biological interaction.

2.2 Explicit Research Gaps

2.2.1 Fragmented pathogen-specific analysis

Many respiratory-virus surveillance assessments focus on individual pathogens or selected pathogen combinations. Such approaches may not adequately represent the broader relationships among multiple circulating viruses. An integrated assessment of ten respiratory virus groups is therefore required to characterize the overall structure of multi-pathogen circulation.

Research gap: Insufficient integration of multiple respiratory virus groups within a unified comparative analytical framework.

Proposed response: Evaluate all ten pathogen groups using common descriptive indicators, pairwise association statistics, correlation matrices, and network representations.

2.22 Limited characterization of pairwise co-detection structures

Identifying the frequency of individual pathogen detections does not explain whether two pathogens tend to be detected together. Conversely, simple correlation analysis does not directly characterize the association between binary detection events.

Research gap: Inadequate joint consideration of binary co-detection relationships and quantitative detection-count associations.

Proposed response: Combine contingency-table analysis, odds ratios, Fisher's exact tests, false-discovery-rate adjustment, and Spearman rank correlations based on log-transformed detection counts.

2.23 Insufficient separation of co-circulation from biological interaction

Country-week co-detection is an ecological observation. It does not establish simultaneous infection in individual patients, direct viral facilitation, competition, or viral interference.

Research gap: Risk of interpreting ecological co-detection and correlation patterns as direct biological interactions.

Proposed response: Establish an explicit inferential distinction between statistical co-circulation, adjusted pathogen associations, and biological interaction. Restrict conclusions from the available dataset to country-week-level observations.

2.24 Limited control of geographical and temporal confounding

Respiratory-virus circulation varies across countries, seasons, and years. Pooled pairwise statistics may therefore capture shared seasonal or geographical patterns rather than pathogen specific associations.

Research gap: Insufficient adjustment for country-level heterogeneity, calendar time, seasonal effects, and other common circulation drivers.

Proposed response: Extend the pairwise analysis through country and time adjusted regression models, stratified analyses, and temporal sensitivity assessments.

2.25 Incomplete treatment of surveillance heterogeneity and missing observations

The dataset exhibits substantial pathogen-level missingness and differences in the number of available observations for individual viruses. A missing pathogen count cannot automatically be interpreted as a zero detection. Moreover, differences in specimen volume and reporting systems may influence observed detection rates.

Research gap: Limited integration of pathogen-specific missingness, testing intensity, and reporting completeness into the interpretation of multi-pathogen associations.

Proposed response: Retain missing values as missing, report pathogen-specific denominators, account for testing intensity where possible, and conduct sensitivity analyses under alternative observation-coverage criteria.

2.26 Limited temporal characterization of multi-pathogen circulation

Monthly detection profiles provide an initial indication of temporal overlap, but descriptive monthly averages cannot fully establish whether pathogen peaks are synchronized, consistently offset, or associated with changes in other pathogens over time.

Research gap: Insufficient investigation of temporal synchronization, lagged associations, and changes in co-circulation patterns.

Proposed response: Extend monthly profiling through seasonality analysis, cross-correlation, lagged association models, and comparisons across defined surveillance periods.

2.27 Incomplete characterization of multi-pathogen circulation burden

Individual pathogen detection rates and pairwise associations do not fully describe the complexity of country-week observations involving several detected pathogens.

Research gap: Limited assessment of the distribution and structure of multi-pathogen detection across country-week observations.

Proposed response: Quantify the number of detected virus groups per eligible country-week and examine how co-circulation burden varies by country, time, and reporting coverage.

2.28 Limited integration of statistical associations into network structures

Pairwise odds ratios and correlation coefficients provide numerical measures of pathogen relationships, but they do not directly reveal the overall connectivity and organization of the multi-pathogen association system.
Research gap: Insufficient integration of pairwise statistical evidence into a unified pathogen-association network.

Proposed response: Construct weighted virus-association networks using statistically supported pairwise associations and evaluate network connectivity, centrality, community structure, and sensitivity to edge-selection criteria.

2.29 Limited robustness and generalizability assessment

Strong associations identified in pooled observations may change when analysis is restricted to particular countries, time periods, or sufficiently reported observations.

Research gap: Insufficient evaluation of the stability of pathogen associations under alternative surveillance and analytical conditions.

Proposed response: Conduct sentinel or higher-coverage subset analyses where source information permits, country-stratified comparisons, alternative detection definitions, and sensitivity analyses using different observation-coverage thresholds.

2.30 Consolidated research-gap statement

The principal research gap is the lack of an integrated analytical approach that simultaneously characterizes pathogen-specific circulation, pairwise co-detection, quantitative detection relationships, temporal overlap, and multi-pathogen burden while explicitly accounting for incomplete surveillance and ecological confounding. The proposed study addresses this gap through a multidimensional analytical framework that combines descriptive epidemiology, statistical association testing, temporal analysis, and pathogen-network modelling. The gaps above are analytical gaps identified from the scope and limitations of the supplied analysis. Their status as gaps in the wider published literature should be established through a systematic literature review before making claims of global research novelty.

3. Research Objectives

3.1 General objective

To develop an integrated statistical and network based framework for characterizing global respiratory virus co-circulation, identifying pairwise pathogen associations, examining temporal circulation patterns, and assessing multi-pathogen detection burden using international weekly surveillance data.

3.2 Specific objectives

RO1. Data quality and surveillance characterization

To assess the geographical coverage, temporal distribution, repeated country week records, pathogen specific observation availability, missingness, and testing-intensity characteristics of the respiratory virus surveillance dataset.

RO2. Pathogen-specific circulation assessment

To quantify and compare the observed detection frequencies, positive country week counts, cumulative reported detections, and positive-observation detection intensities of the ten respiratory virus groups.

RO3. Pairwise co-detection analysis

To identify positive and negative associations between respiratory virus detection indicators using contingency tables, odds ratios, Fisher's exact tests, and Benjamini Hochberg false discovery rate adjustment.

RO4. Quantitative association analysis

To investigate relationships between pathogen detection count magnitudes using Spearman rank correlations of log-transformed reported counts and identify groups of viruses with similar observed circulation patterns.

RO5. Temporal co-circulation analysis

To examine monthly and longer term variations in reported pathogen detection rates, identify periods of overlapping circulation, and investigate potential temporal synchronization or displacement between

pathogens.

RO6. Multi-pathogen burden assessment
To quantify the number of simultaneously detected virus groups within eligible country-week observations and characterize the distribution of single-pathogen and multi-pathogen detection patterns.

RO7. Virus-association network construction
To transform statistically supported pairwise pathogen associations into a weighted network and identify highly connected pathogens, strongly associated pathogen groups, and potential network communities.

RO8. Confounding-adjusted association assessment
To investigate whether the principal pairwise associations remain evident after accounting, where data permit, for geographical heterogeneity, seasonal effects, temporal trends, and surveillance intensity.

RO9. Robustness and sensitivity assessment
To evaluate the stability of the identified co-circulation patterns under alternative observation-coverage thresholds, surveillance subsets, and association definitions.

3.3 Research Questions

The study is designed to address the following research questions:

Research question	**Analytical focus**
RQ1. What are the geographical, temporal, and data-quality characteristics of the global respiratory-virus surveillance dataset?	Coverage, repeated records, missingness, specimen counts, and low-testing observations
RQ2. How do the ten respiratory-virus groups differ in their observed detection frequencies and reported detection intensities?	Pathogen-specific descriptive statistics
RQ3. Which respiratory-virus pairs exhibit the strongest positive or negative country-week co-detection associations?	Odds ratios, contingency tables, Fisher's exact tests, and FDR adjustment
RQ4. Which pathogen pairs demonstrate similar variation in reported detection counts?	Spearman correlation analysis Monthly detection profiles, seasonal analysis, and lagged associations
RQ5. To what extent do respiratory-virus groups exhibit overlapping or differentiated temporal circulation patterns?	Multi-pathogen detection burden

RQ6. How frequently do multiple respiratory-virus groups occur within the same country-week observation?	Weighted association networks,
RQ7. What network structures and pathogen communities emerge from the observed pairwise associations?	centrality, and community detection
RQ8. Do the strongest observed pathogen associations persist after accounting for geographical, temporal, and surveillance-related factors?	Multivariable and stratified association modelling
RQ9. How robust are the identified co-circulation patterns under alternative surveillance-coverage and analytical conditions?	Sensitivity and robustness analysis

3.4 Research Hypotheses

The following hypotheses provide a testable statistical structure for the study. They should be evaluated using the proposed analyses rather than assumed to be true.

H1: Pathogen-specific heterogeneity
The observed detection proportions and detection-count distributions differ across the ten respiratory-virus groups.

- H0: No statistically detectable differences exist in the relevant pathogen-specific detection distributions.

- H1 : At least some pathogen groups exhibit different observed detection distributions.

H2: Pairwise co-detection dependence
The detection indicators of at least some respiratory-virus pairs are statistically associated at the country-week level.

- H0, : Pairwise detection indicators are independent for all evaluated virus pairs.

- H1, : At least one virus pair exhibits a statistically significant departure from independence.

H3: Quantitative detection association
Reported detection counts of some respiratory-virus pairs exhibit positive or negative rank associations.

- H0f : The population rank correlation is zero for the evaluated pathogen pairs.

- H1f : At least one pathogen pair exhibits a non-zero rank correlation.

H4: Temporal heterogeneity

Respiratory-virus groups exhibit differences in their monthly or seasonal detection profiles.

• H04 : The temporal detection profiles do not differ meaningfully among pathogen groups.

• H14 : The temporal detection profiles differ among at least some pathogen groups.

H5: Multi-pathogen detection
Country-week observations with adequate pathogen reporting exhibit multi-pathogen detection patterns rather than exclusively single-pathogen detection.

• H05: The eligible country-week observations do not demonstrate a predominance of multiple detected pathogen groups.

• H15: Multiple pathogen groups are frequently detected within eligible country-week observations.

H6: Persistence of adjusted associations

At least some prominent pairwise associations remain detectable after adjustment for relevant country, temporal, and surveillance-related factors.

• H06 : The identified pairwise associations are not sustained after relevant adjustment.

• H16 : At least some pairwise associations remain statistically supported after adjustment.

H6 is an additional validation hypothesis; it cannot be established from the unadjusted pairwise results alone.

4. Methodological Framework

4.1 Research design

The study adopts a retrospective, observational, quantitative, and ecological research design based on international weekly respiratory virus surveillance data. The analytical unit is the countryweek rather than the individual patient. The framework integrates descriptive statistical analysis, pairwise association testing, temporal analysis, multi-pathogen burden assessment, and network-based characterization.

The proposed methodological structure is organized into nine sequential stages:

1. Dataset acquisition and structural assessment.

2. Data cleaning, aggregation, and quality assessment.

3. Pathogen-specific descriptive analysis.

4. Pairwise co-detection and quantitative association analysis.

5. Temporal circulation and seasonal analysis.

6. Multi-pathogen detection-burden analysis.

7. Weighted virus-association network construction.

8. Confounding-adjusted modelling and sensitivity analysis.

9. Integrated interpretation and epidemiological implications.

4.2 Dataset and pathogen variables

The supplied dataset contains 72,598 raw weekly surveillance records from 184 countries, covering 4 January 2021 to 12 October 2026. Repeated country-week keys occur because multiple reporting-source records can be available for the same country and week. After the exclusion of records with zero or missing specimen counts and consolidation of repeated country-week records, the primary analytical dataset contains 41,465 country-week units.

The ten pathogen groups are:

- Influenza A (INF_A)
- Influenza B (INF_B)
- Adenovirus (ADENO)
- Bocavirus (BOCA)
- Human coronavirus (HUMAN_CORONA)
- Metapneumovirus (METAPNEUMO)
- Parainfluenza (PARAINFLUENZA)
- Rhinovirus (RHINO)
- Respiratory syncytial virus (RSV)
- Other respiratory viruses (OTHERRESPVIRUS)

The analysis should retain available country identifiers, reporting dates, surveillance-source information, specimen counts, low-testing indicators, and relevant geographical and temporal variables.

4.3 Data preprocessing and quality assessment

The first analytical stage establishes the reliability and structure of the surveillance observations.

The preprocessing procedure comprises:

1. Inspection of variable types, unique country identifiers, date ranges, and pathogen-specific variables.

2. Identification of repeated country-week records and assessment of their reporting-source origins.

3. Exclusion of observations with zero or missing specimen counts according to the primary analysis protocol.

4. Consolidation of repeated country-week observations using an aggregation rule appropriate to the underlying reporting structure.

5. Preservation of pathogen-specific missing values rather than automatically replacing them with zero.

6. Identification of country-week observations with low testing intensity.

7. Calculation of pathogen-specific observation coverage and missingness.

8. Verification of specimen denominators and reported detection-count consistency.

The distinction between zero detection and missing observation is fundamental. A zero indicates that a pathogen was observed and no positive detection was reported, whereas a missing value indicates that the pathogen-specific observation is unavailable or unreported. Treating missing observations as zero could generate artificial negative associations and underestimate the extent of circulation.

A further aggregation audit should establish whether repeated records represent additive surveillance counts, separate testing populations, or overlapping reports. Counts should only be summed when the reporting definitions establish that they are non-overlapping and additive. This is necessary to avoid double-counting specimens or pathogen detections.

4.31 Pathogen-specific descriptive analysis

The second stage quantifies the observed circulation characteristics of each pathogen.

For pathogen (i), the positive detection proportion is defined as:

$$P_i = \frac{N_i^+}{N_i^{obs}}$$

- ($N+_i$represents the number of country-week observations with a positive reported detection for pathogen i.
- $Nobs_i$ represents the number of country-week observations with non-missing pathogen-specific data.
- The total reported detections are calculated as:

$$D_i = \sum_{t=1}^{N_i^{obs}} X_{it}$$

- where X_{it} represents the reported detection count for pathogen i in observation t.
- The analysis should additionally report:
- Observed country-week count.
- Positive country-week count.
- Positive detection proportion.
- Total reported detections.
- Median and interquartile range of positive counts.
- Detection-rate distributions where valid specimen denominators are available.

These statistics establish the descriptive baseline for subsequent association and temporal analyses. Since pathogen-specific denominators differ, detection proportions should not be interpreted as directly comparable population prevalence estimates.

4.32 Pairwise co-detection analysis

$$Z_{it} = \begin{cases} 1, & X_{it} > 0 \\ 0, & X_{it} = 0 \end{cases}$$

The binary variable is defined only where a valid pathogen-specific observation is available. Missing values remain missing.

For every pathogen pair (*i, j*), a 2x2 contingency table is constructed:

	Pathogen *j* Positive	Pathogen *j* negative
Pathogen *i* Positive	*a*	*b*
Pathogen *i* negative	*c*	*d*

The odds ratio is estimated as:

$$OR_{ij} = \frac{ad}{bc}$$

where a, b, c and d denote the corresponding contingency-table frequencies.

Interpretation:

- OR_{ij} >1: positive statistical co-detection association.
- OR_{ij} <1: negative statistical co-detection association.
- OR_{ij} =1: no association under the odds-ratio measure.

Fisher's exact test is used to assess the statistical significance of the pairwise association. Since ten pathogen groups generate 45 unique pathogen pairs, the Benjamini–Hochberg procedure is applied to control the false discovery rate.

The analysis should report the complete-pair denominator, contingency-table frequencies, co-positive observations, odds ratios, confidence intervals, raw p-values, and adjusted q-values.

Confidence intervals are important because an odds ratio alone does not communicate the precision of the estimated association. Where contingency cells are zero, an appropriate exact or penalized estimation procedure should be used rather than reporting an undefined or infinite estimate without qualification.

4.4 Quantitative pathogen association analysis

Binary co-detection analysis identifies whether two pathogens are detected together, but it does not capture the relationship between their reported detection magnitudes.

To complement the binary analysis, Spearman rank correlations are estimated using log-transformed reported counts:

$$Y_{it} = \log(1 + X_{it})$$

The Spearman coefficient is:

$$\rho_{ij} = \mathbf{Corr}\left[\mathbf{Rank}'(Y_i)\mathbf{Rank}(Y_j)\right]$$

where the correlation is calculated using country-week observations with valid measurements for both pathogens.

The analysis produces a ten by ten correlation matrix, a correlation heatmap, and a ranked list of pathogen pairs.

Positive correlations indicate that higher reported detection counts of one pathogen tend to coincide with higher counts of another. Negative correlations indicate an inverse association. These relationships remain observational and may reflect common temporal or geographical drivers.

4.41 Temporal co-circulation and seasonal analysis

The fourth major analytical component investigates the temporal variation in pathogen circulation.

The existing analysis calculates monthly mean reported detections per 1,000 specimens. The descriptive rate for pathogen (i) in country-week observation (t) can be expressed as:

$$R_{it} = \frac{X_{it}}{S_t} \times 1000$$

where:

- X_{it} is the reported detection count.
- S_t is the corresponding specimen count.

This formulation is valid only when the numerator and denominator represent compatible reporting populations and measurement definitions. If the underlying surveillance data do not establish this compatibility, the resulting rate should be described as a descriptive specimen-normalized indicator rather than a standardized positivity estimate.

Monthly profiles are then generated to identify:

- Periods of elevated pathogen detection.
- Overlapping monthly peaks.
- Differences in seasonal timing.
- Potential temporal displacement.
- Changes in observed circulation across the surveillance period.

The proposed extension should investigate temporal dependence using cross correlation and lagged association models. For example:

$$C_{ij}(k) = \mathrm{Corr}\left(Y_{i,t'}\, Y_{j,t-k}\right)$$

where k represents the temporal lag.

The analysis can evaluate selected lags, such as one to twelve weeks, where the temporal coverage and reporting frequency support such comparisons. Multiple-lag testing should be accounted for, and country-specific temporal profiles should be examined before interpreting pooled results.

For a stronger seasonal assessment, monthly observations can be analysed using seasonal decomposition or other time-series approaches, provided that missing periods and irregular reporting are handled appropriately.

4.42 Multi-pathogen detection-burden analysis

The fifth stage quantifies the number of virus groups detected within each country-week observation.

For a country-week t, the observed number of detected pathogen groups is:

$$B_t = \sum_{i=1}^{10} Z_{it}$$

where Z_{it} represents the binary detection indicator.

However, this measure is interpretable only when the number of pathogen groups with valid observations is considered. The primary analysis restricts the burden assessment to country-weeks with at least five pathogen groups reported.

The study should report:

- Number of country-weeks with zero detected groups.
- Number with one detected group.
- Number with two or more detected groups.
- Distribution of observations across detection-burden categories.
- Proportion of eligible observations with multiple detections.

Variation in burden across countries and time periods.

A normalized burden measure may also be considered:

$$B_t^* = \frac{\sum_i z_{it}}{M_t}$$

- where Mt represents the number of pathogen groups with non-missing observations in country-week t.

This supplementary measure helps distinguish the number of detected pathogens from the proportion of observed pathogen groups that are positive. Neither measure establishes individual-level coinfection.

4.5 Weighted virus-association network

The sixth stage transforms pairwise pathogen relationships into a network representation.

The network is defined as:

$$G = (V, E, W)$$

where:

- V represents the ten respiratory-virus groups.
- E represents the statistically supported associations between pathogen pairs.
- W represents the corresponding association weights.

Each pathogen is represented as a node, and an edge connects two pathogens when the selected statistical criterion is satisfied.

Two complementary network representations are recommended:

Network A: Binary co-detection network

Edges are based on pairwise odds ratios. Edge weights may be defined using the logarithm of the odds ratio:

$$w_{ij}^{OR} = \log(OR_{ij})$$

This representation distinguishes positive and negative associations through the sign of the weight. It should include only estimates with adequate precision and clearly specified statistical support.

NetworkB:Detection-countcorrelationnetwork

Edges are weighted using Spearman correlation coefficients:

$$w_{ij}^{\rho} = \rho_{ij}$$

This network represents similarities in quantitative detection patterns.

The following network measures can be examined:

Network measure	**Purpose**
Degree	Number of statistically supported connections associated with each pathogen
Weighted degree / strength	Total magnitude of connections associated with a pathogen
Betweenness centrality	Extent to which a pathogen lies on shortest paths between other nodes
Closeness centrality	Relative network proximity of a pathogen to other pathogens
Eigenvector centrality	Connectivity to other highly connected pathogens
Community detection	Identification of groups of pathogens with comparatively strong internal associations
Network density	Overall proportion of possible pathogen connections represented in the network

With only ten pathogen nodes, network statistics should be interpreted as descriptive summaries rather than definitive evidence of epidemiological importance. Community detection and centrality estimates should be checked for sensitivity to the edge-selection threshold and association measure.

Importantly, the network represents statistical relationships among surveillance detections, not a transmission network or a network of proven biological interactions.

4.51 Confounding-adjusted association modelling

A central extension beyond the current descriptive and unadjusted pairwise analyses is to assess whether the strongest associations remain after accounting for shared country, time, and surveillance characteristics. For each pathogen pair, a logistic regression model may be specified as:

$$\mathbf{logit}\big(P(Z_{it} = 1)\big) = \beta_0 + \beta_1 Z_{jt} + \gamma^T C_t + \delta^T T_t + u_c$$

where:

- Z_{it} is the detection indicator for the outcome pathogen.
- Z_{jt} is the detection indicator for the paired pathogen.
- C_t represents available surveillance and testing covariates.
- T_t represents temporal variables, such as month, year, and seasonal indicators.
- u_c represents a country-level random intercept where a mixed-effects model is appropriate.

The adjusted odds ratio is:

$$AOR_{ij} = \exp(\beta_1)$$

The model should incorporate suitable country-level dependence structures and, where justified, temporal autocorrelation. Candidate surveillance covariates may include specimen volume, low-testing status, reporting source, and available measures of observation coverage.

A country fixed-effects approach may be considered when the objective is to control for stable country-specific characteristics. A mixed-effects approach may be more appropriate when estimating between-country variability. Model choice should be guided by the data structure, number of observations per country, and convergence diagnostics.

Where the available data do not contain compatible testing-population information or sufficient observations for a particular pathogen pair, adjusted estimates should be reported as unavailable rather than inferred.

The adjusted analysis should compare the direction, magnitude, and uncertainty of the associations with the original unadjusted odds ratios. A substantial reduction in an association after adjustment would indicate that common temporal, geographical, or surveillance characteristics may explain part of the pooled relationship.

4.52 Robustness and sensitivity analysis

The final analytical stage assesses the stability of the principal findings.

Recommended sensitivity analyses include:

1. Observation-coverage thresholds: Repeat the co-detection burden analysis using alternative minimum numbers of observed pathogen groups.

2. Surveillance-source stratification: Compare sentinel and non-sentinel observations where source classifications and sample sizes support valid comparisons.

3. Low-testing exclusion: Repeat selected analyses after excluding observations flagged as low testing.

4. Country-stratified analysis: Examine whether major associations are consistent across countries with sufficient observations.

5.Temporal stratification: Compare selected periods within the surveillance window, using clearly defined period boundaries.

6. Alternative association measures: Compare odds ratios, phi coefficients, and Spearman correlations.

7. Confidence-interval assessment: Evaluate whether strong pairwise associations are supported by sufficiently precise estimates.

8. Missingness assessment: Compare results under alternative defensible observation-availability criteria without automatically imputing missing pathogen counts as zero.

9. Network stability: Reconstruct networks using alternative statistical thresholds and association weights.

10. Lagged sensitivity: Compare contemporaneous and lagged associations to determine whether relationships are dependent on the selected temporal window.

The sensitivity analysis should identify associations that are consistently supported and distinguish them from relationships that are highly dependent on particular surveillance or modelling assumptions.

5. Proposed Integrated Methodological Architecture

The overall research architecture can be represented as the following sequence:

Input layer: International weekly respiratory-virus surveillance dataset.
Stage 1 – Data quality and preprocessing: Country-week identification, duplicate-source assessment, specimen-count validation, missingness analysis, and testing-intensity assessment.

Stage 2 – Pathogen circulation characterization: Detection frequency, positive observations, cumulative detections, and pathogen-specific descriptive statistics.

Stage 3 – Pairwise statistical analysis: Binary co-detection, odds ratios, Fisher's exact tests, FDR correction, and Spearman correlations.

Stage 4 – Temporal analysis: Monthly detection rates, seasonal profiles, temporal overlap, and lagged association assessment.

Stage 5 – Multi-pathogen burden: Number and proportion of detected virus groups within sufficiently observed country-week units.

Stage 6 – Network construction: Weighted pathogen-association networks, centrality measures, and community identification.

Stage 7 – Adjusted inference and validation: Country- and time-adjusted modelling, surveillance sensitivity analysis, and network robustness assessment.

Output layer: A validated descriptive and statistical representation of global respiratory virus co-circulation, together with an evidence based interpretation of pathogen association patterns and surveillance limitations.

5.1 Expected Research Contributions

The proposed study is expected to provide the following contributions:

Theoretical contribution: It frames respiratory-virus circulation as a multi-pathogen ecological system rather than a collection of independent pathogen-specific processes.

Methodological contribution: It integrates binary co-detection statistics, quantitative detection correlations, temporal analysis, burden assessment, and network modelling within a common analytical framework.

Empirical contribution: It provides an international, multi-year characterization of ten respiratory-virus groups across 184 countries using the supplied surveillance dataset.

Epidemiological contribution: It identifies pathogen combinations and temporal periods that may warrant more detailed investigation, while distinguishing descriptive association from biological interaction.

Surveillance contribution: It highlights the importance of pathogen-specific reporting completeness, specimen denominators, and testing intensity when interpreting global respiratory-virus surveillance data.

Practical contribution: It provides an analytical foundation for integrated respiratory-virus monitoring and for prioritizing pathogen combinations for further epidemiological or clinical investigation.

5.2 Methodological Limitations and Interpretation Boundaries

Several limitations must be explicitly recognized.

First, the country-week unit is ecological. It cannot establish that two pathogens were detected in the same individual or that one pathogen directly influenced the infection dynamics of another.

Second, pathogen-specific missingness is substantial. A missing result cannot be treated as a negative result, and comparisons across pathogens must account for unequal observation denominators.

Third, the pooled odds ratios and correlations may be influenced by shared seasonality, country characteristics, surveillance practices, testing intensity, and changes in reporting systems.

Fourth, the descriptive detection rate per 1,000 specimens is not necessarily equivalent to a pathogen-specific positivity rate unless the numerator and denominator refer to compatible tested populations.

Fifth, temporal correlations and network structures can be sensitive to the choice of aggregation period, edge threshold, and missing-data criteria.

Finally, the proposed adjusted and lagged analyses require adequate country-week coverage and relevant covariates. Their feasibility must be verified before they are presented as completed results.

Accordingly, the study should use the terms co-circulation, co-detection, and statistical association when describing the existing findings. Terms such as viral interference, facilitation, or biological interaction should be reserved for conclusions supported by suitable individual-level, experimental, or otherwise appropriately controlled evidence.

10. Overall Research Framework and Expected Outcome

The proposed research establishes a comprehensive framework for investigating global respiratory-virus co-circulation through the integration of surveillance-data assessment, pathogen specific descriptive statistics,

pairwise association testing, temporal profiling, multi-pathogen burden analysis, and weighted network modelling. Its central methodological emphasis is the identification of statistically meaningful pathogen relationships while accounting for the limitations imposed by incomplete and heterogeneous surveillance reporting.

The framework moves from descriptive characterization toward increasingly structured statistical inference. The initial analysis establishes which pathogens are frequently detected and which combinations exhibit elevated co-detection. Temporal and burden analyses then characterize when and how multiple pathogens are reported together. Network analysis provides a system-level representation of these relationships, while adjusted modelling and sensitivity assessments determine whether the principal associations remain robust after accounting for shared geographical, temporal, and surveillance-related influences.

The expected outcome is an integrated and reproducible analytical representation of respiratory-virus circulation that identifies prominent pathogen associations, characterizes temporal overlap, quantifies multi-pathogen detection burden, and distinguishes robust statistical relationships from potentially confounded ecological patterns. This provides a defensible foundation for subsequent epidemiological investigation without overstating the biological meaning of surveillance-based co-detection.

6. Main results

6.1 Virus-specific circulation

Table 2 summarizes the number of country-weeks with available pathogen observations and the proportion of those observations with a positive reported count. The denominator differs by virus because pathogen-specific reporting is incomplete; therefore, proportions should be compared cautiously.

6.2 Pairwise co-detection and association

The largest estimated co-detection odds ratio in the pairwise screen was for ADENO and METAPNEUMO (OR=20.33; 4,387 co-positive country-weeks among 11,292 complete pairs). This indicates that joint detection was more frequent than expected under independence in the pooled country-week contingency table. It does not show that either virus biologically facilitates the other.

Odds ratios are unadjusted and can be influenced by shared seasonality, geography, surveillance intensity, testing policy, and differences in reporting completeness. Fisher exact test p-values were adjusted across the 45 virus pairs using the Benjamini Hochberg false discovery rate procedure. Spearman correlations use log1p-transformed reported counts and are likewise observational.

6.3 Temporal circulation

Figure 4 presents monthly mean reported detections per 1,000 specimens, calculated within country-week observations and then averaged by month and virus. This is a descriptive normalization only: specimen denominators may not represent the same pathogen-specific testing population across countries or reporting systems.

6.4 Co-circulation burden

The number of simultaneously detected virus groups was counted for country-weeks with at least five virus groups reported. This reduces the influence of very sparse pathogen reporting but does not eliminate surveillance bias.

Documents the raw sample size, country coverage, duplicated country-week keys, retained specimen-positive records, and pathogen missingness. High missingness means the absence of a reported value cannot be interpreted as the absence of circulation.

Metric	Value	Interpretation
Raw weekly records	72598	Records in source file
Countries	184	Unique country names
Date range	2 021-01-04 to 2026-10-12	Weekly surveillance coverage
Country-week records after aggregation	41465	Unique country-week units
Raw records with specimens > 0	55742	Retained for primary analysis
Total specimens reported	64677119	Sum of reported specimen counts
Duplicate country-week keys in raw data	23739	Multiple reporting-source records can occur
Mean missingness across virus variables	0.618	Missing pathogen counts are not treated as zero
Country-weeks flagged low testing	7805	Quality flag carried into aggregated data

Table 1. Data quality and analytical coverage

Documents the raw sample size, country coverage, duplicated country-week keys, retained specimen-positive records, and pathogen missingness. High missingness means the absence of a reported value cannot be interpreted as the absence of circulation.

Table 1 presents the data quality assessment and overall analytical coverage of the respiratory-virus surveillance dataset. The original dataset comprised 72,598 weekly surveillance records collected from 184 countries between 4 January 2021 and 12 October 2026. Following the exclusion of records with zero or missing specimen counts and the consolidation of repeated country-week observations, 41,465 unique country-week units were retained for subsequent analyses. The original dataset contained 23,739 repeated country-week keys, indicating that multiple reporting records were available for certain country-week combinations.

A total of 55,742 records contained positive specimen counts, representing the observations eligible for the primary analysis. The cumulative number of reported specimens was 64,677,119. Importantly, the mean missingness across the ten respiratory-virus variables was 61.8%, indicating substantial incompleteness in pathogen-specific surveillance reporting. Furthermore, 7,805 country-week observations were flagged as having low testing intensity.

These findings demonstrate that the dataset provides extensive geographical and temporal coverage but also contains considerable heterogeneity in reporting completeness. The high proportion of missing pathogen observations is particularly important because an unreported viral count cannot automatically be interpreted as the absence of viral circulation. Consequently, retaining missing values rather than replacing them with zeros is essential to avoid underestimating virus prevalence and generating artificial negative associations between pathogens. The presence of low-testing observations also suggests that surveillance intensity should be considered when interpreting differences in virus detection across countries and time periods.

The data quality assessment establishes the reliability boundaries of the subsequent co-circulation analyses and highlights the need for missingness-aware statistical modelling and sensitivity analysis.

Table 2. Virus-specific circulation summary
Compares the availability of pathogen observations, positive country-weeks, positive proportions, total reported detections, and median counts among positive observations. Denominators vary across viruses.

Virus-Observed	country weeks	Positive country weeks	Positive proportion among observed	Total reported detections	Median count when positive
INF_A	33108	25986	0.784885828	4 395375	10
INF_B	36391	18211	0.500425929	970773	5
ADENO	12219	6544	0.535559375	130737	5
BOCA	9341	2297	0.245905149	7754	2
HUMAN_CORONA	9685	3791	0.391430046	92905	4
METAPNEUMO	11695	5251	0.448995297	161564	6
PARAINFLUENZA	11306	5747	0.508314169	192367	6
RHINO	10952	6060	0.553323594	491673	13
RSV	27834	17678	0.635122512	1131121	8
OTHERRESPVIRUS	11234	3667	0.326419797	112470	6

Table 2. Descriptive statistics of respiratory-virus circulation

Table 2 summarizes the circulation characteristics of ten respiratory-virus groups using the number of observed country-weeks, positive country-weeks, positive detection proportions, total reported detections, and median detection counts among positive observations.

Influenza A (INF_A) exhibited the highest observed detection proportion, with 25,986 positive country-weeks among 33,108 observations (78.49%). RSV was the second most frequently detected pathogen, with 17,678 positive observations among 27,834 country-weeks (63.51%). Rhinovirus (RHINO) and adenovirus (ADENO) also demonstrated relatively high detection proportions of 55.33% and 53.56%, respectively. Parainfluenza and influenza B showed intermediate detection proportions of 50.83% and 50.04%, respectively.

In contrast, bocavirus (BOCA) exhibited the lowest detection proportion, with 2,297 positive observations among 9,341 country-weeks (24.59%). Other respiratory viruses and human coronavirus also demonstrated comparatively lower detection proportions of 32.64% and 39.14%, respectively. Regarding detection intensity, influenza A recorded the largest cumulative number of detections (4,395,375), followed by RSV (1,131,121) and influenza B (970,773). Rhinovirus exhibited the highest median count among positive observations, with

a median of 13, followed by influenza A, with a median of 10.

Overall, the results indicate considerable variation in the geographical and temporal distribution of respiratory viruses. Influenza A and RSV were prominent in terms of observed detection frequency, whereas rhinovirus demonstrated comparatively high detection intensity among positive observations. The differences between cumulative detections and median positive counts also indicate that overall circulation burden and typical detection intensity represent distinct epidemiological characteristics.

However, the pathogen specific observation denominators differ substantially because of incomplete surveillance reporting. Therefore, the observed detection proportions should not be interpreted as directly comparable population level prevalence estimates.

Table 3 presents the pairwise statistical assessment of respiratory virus co-circulation. The analysis evaluates combinations of the ten pathogen groups using complete country week observations for each virus pair. The reported indicators include the number of country-weeks with observations for both viruses, the number of simultaneously positive observations, co-positive proportions, odds ratios, Fisher's exact-test p-values, false-discovery rate adjusted q-values, Spearman correlation coefficients, and the corresponding contingency-table frequencies.

The odds ratio provides an assessment of whether the simultaneous detection of two viruses occurs more frequently or less frequently than expected under statistical independence. An odds ratio greater than one indicates a positive association between the two detection indicators, whereas an odds ratio below one indicates a negative association.

The strongest reported co-detection association was observed between adenovirus and metapneumovirus (OR = 20.33), with 4,387 co-positive country weeks among 11,292 complete observations. This result indicates a substantially higher frequency of simultaneous detection than would be expected under independence within the pooled surveillance observations.

The statistical associations were further evaluated using Fisher's exact test, with Benjamini–Hochberg adjustment applied across the 45 virus pairs to control the false discovery rate. This adjustment is important because multiple pairwise comparisons increase the probability of identifying apparently significant associations by chance.

The Spearman correlation analysis complements the binary co-detection analysis by examining relationships between the magnitudes of reported viral detections. The correlations were calculated using log-transformed detection counts, allowing the analysis to examine whether higher reported counts of one pathogen tend to coincide with higher counts of another pathogen.

Taken together, the results provide evidence of heterogeneous relationships among respiratory-virus groups. Nevertheless, the observed associations should be interpreted as ecological statistical relationships rather than evidence of direct viral facilitation, competition, or interference. Shared seasonal patterns, geographical characteristics, testing intensity, and differences in surveillance practices may contribute to the observed associations.

Pairwise co-detection analysis identifies potentially important pathogen combinations and provides a basis for more advanced multivariable and temporal interaction modelling.

Reports complete-pair denominators, co-positive country-weeks, odds ratios, Fisher exact p-values, FDR-adjusted q-values, Spearman correlations, and the underlying 2×2 cell counts. Interpret OR>1 as positive co-detection association and OR<1 as negative association, not as proof of biological interaction.

Number of viruses detected	Country-weeks	Proportion
0	946	0.076038904
1	1563	0.125632988
2	1805	0.1450848
3	1160	0.093240093
4	980	0.078771803
5	1104	0.088738847
6	1278	0.102724861
7	998	0.080218632
8	1367	0.109878627
9	901	0.072421831
10	339	0.027248613

Table 3. Pairwise co-circulation and interaction statistics

Table 4 presents the monthly detection-rate characteristics of the ten respiratory-virus groups. The detection rates were calculated by dividing the reported viral detections by the corresponding number of specimens and expressing the resulting values per 1,000 specimens. Monthly averages were subsequently calculated using the available country-week observations for each pathogen.

This table provides a temporal perspective on respiratory-virus circulation by allowing the comparison of detection-rate patterns across different months and pathogens. The inclusion of the number of contributing country-week observations is particularly important because it indicates the surveillance coverage underlying each monthly estimate.

Differences in monthly detection rates can help identify periods of increased or reduced viral circulation. Similarly, overlapping increases across multiple pathogens may indicate periods of concurrent circulation, whereas differences in the timing of monthly peaks may suggest distinct seasonal circulation patterns.

The normalization by specimen counts provides a descriptive measure of detection intensity relative to reported testing volume. However, it does not completely eliminate differences in surveillance practices because the specimen denominators may not represent identical pathogen-specific testing populations across countries and reporting systems.

Scientific implication: Monthly detection rate patterns provide an epidemiological basis for identifying seasonal overlap, potential temporal displacement, and periods of intensified multi-virus circulation.

Provides month by month mean reported detections per 1,000 specimens for each virus, along with the number of country week observations contributing to each estimate.

Shows how frequently zero, one, or multiple virus groups were detected in country-weeks with at least five pathogen groups reported.

Number of viruses detected	Country-weeks	Proportion
0	946	0.076038904
1	1563	0.125632988
2	1805	0.1450848
3	1160	0.093240093
4	980	0.078771803
5	1104	0.088738847
6	1278	0.102724861
7	998	0.080218632
8	1367	0.109878627
9	901	0.072421831
10	339	0.027248613

Table 5. Co-circulation burden

Table 5 describes the distribution of the number of respiratory virus groups detected simultaneously within individual country week observations. To reduce the influence of observations with extremely limited pathogen reporting, the analysis was restricted to country weeks in which at least five virus groups had reported observations.

The results demonstrate substantial variation in the number of simultaneously detected pathogens. A total of 946 country weeks (7.60%) had no detected viruses, while 1,563 country weeks (12.56%) recorded one detected virus. Two and three viruses were detected in 1,805 (14.51%) and 1,160 (9.32%) country weeks, respectively.

Higher co-circulation burdens were also observed. In particular, 1,367 country weeks (10.99%) recorded eight simultaneously detected virus groups, while 901 (7.24%) recorded nine detected groups. The detection of all ten virus groups occurred in 339 country weeks, corresponding to 2.72% of the observations.

Collectively, 9,932 of the 12,441 eligible country weeks (79.83%) recorded at least two detected virus groups. This finding indicates that multi-virus circulation was a prominent feature of the observed surveillance dataset. Nevertheless, these results describe the simultaneous presence of detected pathogen groups at the country-week level and do not establish that multiple viruses infected the same individuals.

Scientific implication: The co-circulation burden analysis demonstrates the extent of simultaneous respiratory-virus activity and provides a quantitative basis for investigating complex multi-pathogen circulation patterns. Spearman Matrix- Virus Co-circulation Table

	INF_A	INF_B	ADENO	BOCA	HUMAN_CORONA	METAP NEUMO	PARAIN FLUENZA
Inf_A	1	0.584354877	0.529564166	0.29375783	0.405852473	0.560946279	0.494433738
Inf_B	0.584354877	1	0.491513829	0.239598979	0.38506523	0.500683101	0.435908716
Adeno	0.529564166	0.491513829	1	0.42977524	0.639889539	0.752684167	0.777012144
Boca	0.29375783	0.239598979	0.42977524	1	0.479332584	0.424799474	0.458672998
Human_Corona	0.405852473	0.38506523	0.639889539	0.479332584	1	0.589189377	0.654579788
Metapneumo	0.560946279	0.500683101	0.752684167	0.424799474	0.589189377	1	0.73852815
Parainfluenza	0.494433738	0.435908716	0.777012144	0.458672998	0.654579788	0.73852815	1
Rhino	0.436882642	0.372048598	0.701835599	0.451073246	0.59110868	0.691929729	0.679847351
Rsv	0.558283559	0.449030587	0.641706008	0.338646303	0.46143608	0.65469221	0.612748431
Otherrespvirus	0.383190845	0.284614827	0.571148	0.397001281	0.542162675	0.566326884	0.599597713

Table 6. Spearman correlation matrix of respiratory-virus co-circulation

The Spearman correlation matrix examines the relationships between reported detection counts across the ten respiratory virus groups. The analysis uses log-transformed detection counts and provides a complementary assessment of pathogen co-circulation beyond the binary detection indicators presented in Table 3.

The results demonstrate predominantly positive correlations between respiratory-virus detection counts, indicating that higher reported detection levels of certain pathogens tend to coincide with higher detection levels of other pathogens across the observed country-week units.

The strongest correlation was observed between adenovirus and parainfluenza (ρ=0.777), followed by adenovirus and metapneumovirus (ρ= 0.753). A similarly strong positive relationship was observed between metapneumovirus and parainfluenza (ρ= 0.739). These associations suggest that these pathogens exhibit relatively similar patterns of reported detection intensity across the surveillance observations.

Other notable correlations included adenovirus and rhinovirus (ρ= 0.702), metapneumovirus and rhinovirus (ρ= 0.692), and parainfluenza and rhinovirus (ρ= 0.680). In contrast, weaker relationships were observed for bocavirus and influenza B (ρ = 0.240), as well as bocavirus and influenza A (ρ= 0.294).

The observed correlation structure suggests that some respiratory viruses exhibit stronger similarities in their geographical and temporal detection patterns than others. In particular, adenovirus, metapneumovirus, parainfluenza, and rhinovirus form a group of relatively strongly correlated pathogens.

However, positive correlations do not necessarily indicate direct biological interactions. These relationships may arise from common seasonal patterns, shared geographical distributions, changes in testing intensity, or similar surveillance coverage.

The correlation matrix identifies groups of pathogens with similar circulation patterns and provides a foundation for constructing virus-association networks and investigating potential common seasonal drivers.

6.5 Further analytical interpretations

Figure 1. Observed frequency of respiratory-virus circulation

Ranks virus groups by the proportion of non-missing country-week observations with a positive reported count. It identifies frequently detected pathogens within their available reporting windows; unequal denominators limit direct ranking across pathogens.

Figure 1. Observed frequency of respiratory-virus circulation
Description and interpretation

Figure 1 illustrates the relative frequency of detection of the ten respiratory-virus groups across the available country-week surveillance observations. The figure ranks pathogens according to the proportion of non-missing country-week observations in which at least one positive detection was reported.

Influenza A exhibits the highest observed detection frequency, followed by RSV, rhinovirus, and adenovirus. These pathogens therefore represent prominent components of the respiratory-virus circulation patterns captured by the surveillance dataset. Conversely, bocavirus exhibits the lowest observed detection frequency, indicating that positive detections were less frequently reported for this pathogen within its available surveillance observations.

The differences in detection frequency suggest that respiratory-virus circulation is heterogeneous across pathogens. However, the observed proportions reflect surveillance detections rather than standardized population-level infection prevalence. Differences in testing coverage, geographical representation, and reporting completeness may influence the relative ranking.

Figure 1 establishes the relative detection frequency of respiratory pathogens and identifies the viruses that contribute most prominently to the observed surveillance burden.

Displays rank correlations between log-transformed pathogen detection counts for country-weeks with paired observations. Positive values indicate that larger counts tend to co-occur; common seasonality and specimen volume may explain part of these associations.

Figure 2 presents a correlation heatmap illustrating the pairwise Spearman relationships between the log-transformed detection counts of the ten respiratory-virus groups. The colour intensity represents the magnitude and direction of the correlation coefficients, with positive correlations indicating that higher detection counts of one pathogen tend to occur alongside higher detection counts of another pathogen.

The heatmap reveals predominantly positive relationships among the respiratory-virus groups. Particularly strong correlations are visible between adenovirus and parainfluenza, adenovirus and metapneumovirus, and metapneumovirus and parainfluenza. The relationships between adenovirus and rhinovirus and between metapneumovirus and rhinovirus are also comparatively strong.

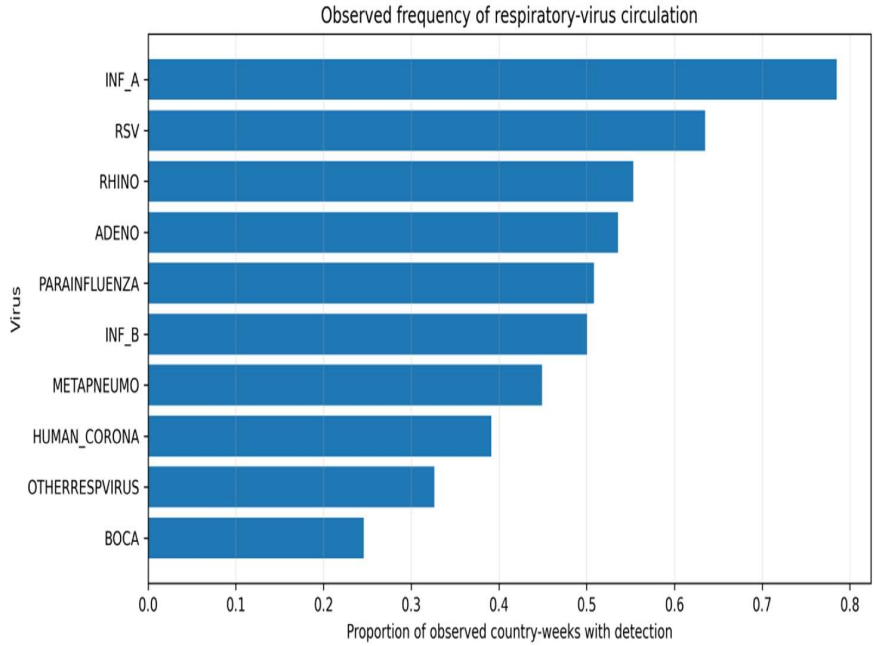

Figure 2. Pairwise Spearman correlation heatmap

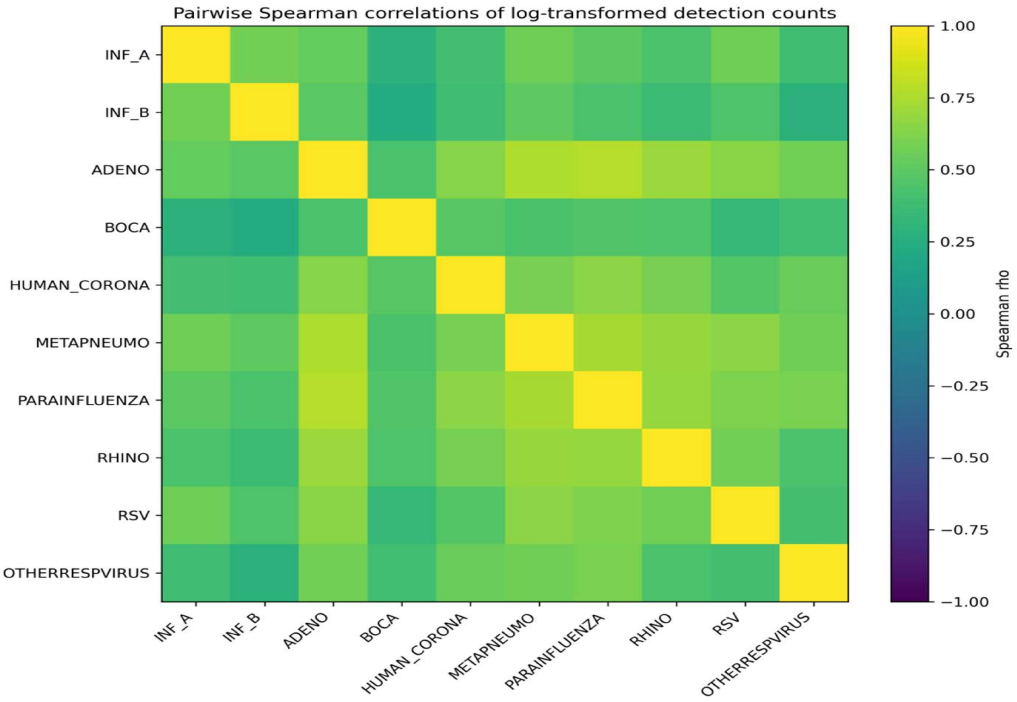

Figure 2A. Pairwise Spearman correlation heatmap of respiratory-virus detection counts

These patterns indicate that several respiratory pathogens exhibit similar reported circulation intensities across the surveillance observations. Such similarities may reflect shared seasonal activity, geographical overlap, or comparable surveillance patterns. In contrast, the relatively weak correlations involving bocaviru s and influenza B suggest that their reported detection counts exhibit less consistent variation together.

The overall correlation structure highlights the presence of pathogen groups with similar circulation profiles. Nevertheless, the heatmap should be interpreted as an exploratory representation of statistical associations rather than evidence of direct biological interaction or causal relationships.

The correlation heatmap identifies strongly associated pathogen pairs and provides an analytical basis for identifying potential clusters of respiratory viruses with similar circulation dynamics.

Figure 3. Pairwise co-detection odds-ratio heatmap

Shows the strength of pairwise association between binary detection indicators. Values above one indicate positive co-detection association and values below one indicate negative association. The logarithmic scale makes relative differences easier to compare.

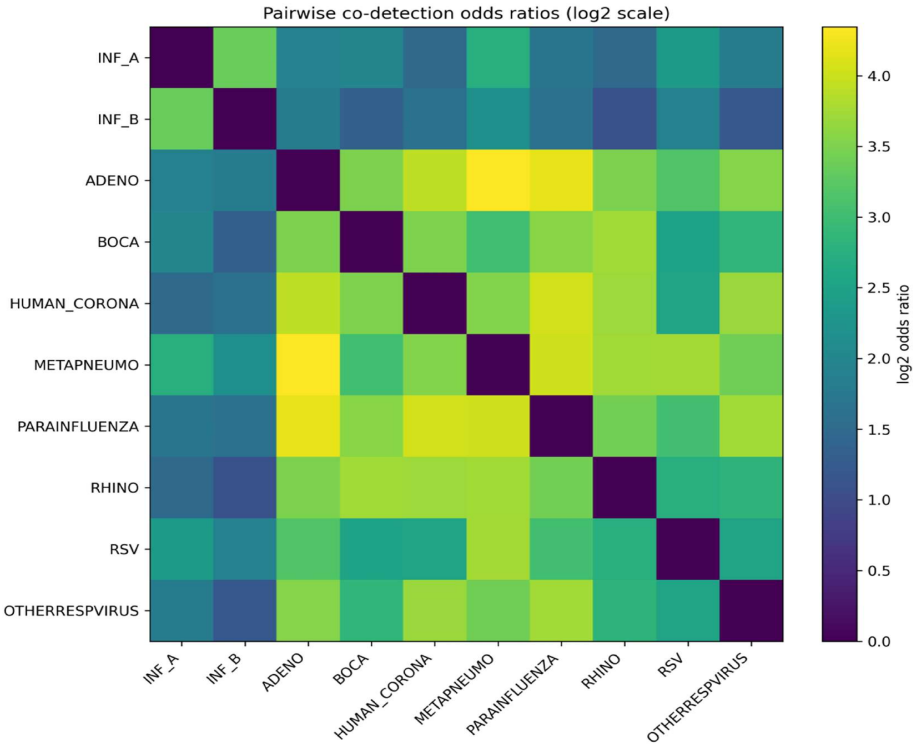

Figure 3. Pairwise co-detection odds-ratio heatmap

Figure 3 visualizes the pairwise odds ratios calculated from binary virus-detection indicators. The heatmap provides a comparative representation of whether two respiratory viruses are detected together more or less frequently than expected under statistical independence.

The logarithmic transformation of the odds ratios improves the visualization of differences in association strength across pathogen pairs. Positive log transformed values correspond to odds ratios greater than one, indicating positive co-detection associations, whereas negative values correspond to odds ratios below one, indicating negative associations. Values close to zero represent odds ratios close to one.

The prominent association between adenovirus and metapneumovirus is particularly noteworthy, consistent with the highest reported pairwise odds ratio of 20.33. This indicates a strong positive statistical association between their simultaneous detection indicators. The presence of other elevated pairwise associations suggests that respiratory-virus co-detection is not uniformly distributed across pathogen combinations.

The heatmap therefore provides an overview of the statistical structure of respiratory-virus co-circulation and identifies combinations that may warrant further epidemiological investigation. However, the magnitude of an odds ratio does not independently establish biological interaction. Shared seasonal circulation, country-specific characteristics, and differences in surveillance coverage can produce strong associations even in the absence of direct viral interactions.

Scientific implication: Figure 3 identifies potentially important pathogen combinations and helps distinguish positive from negative co-detection relationships, providing a foundation for more rigorous interaction modelling.

Shows monthly mean detections per 1,000 specimens for each pathogen. Overlapping peaks suggest temporal co-circulation, whereas offset peaks suggest different seasonal timing.

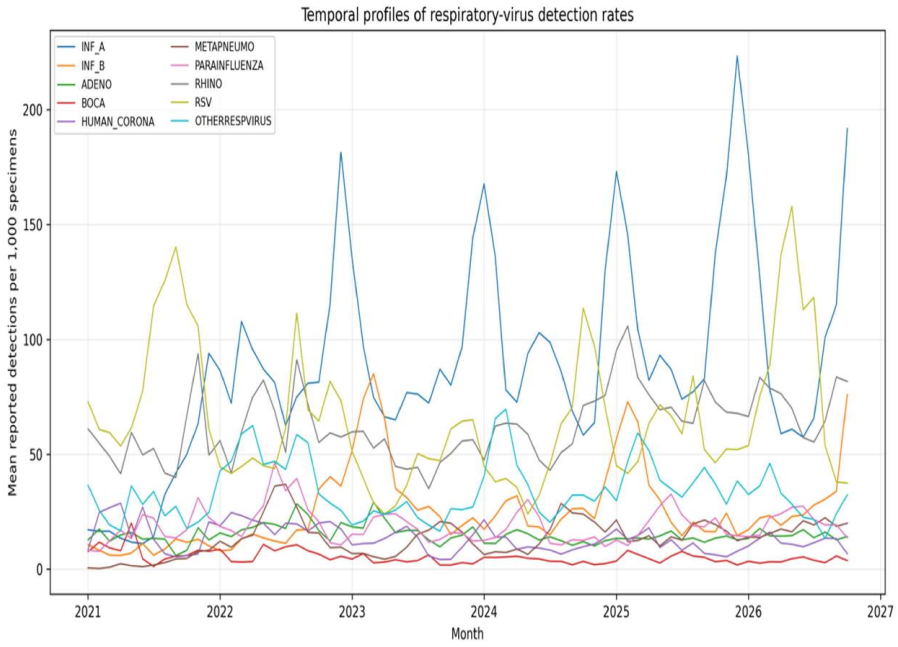

Figure 4. Temporal profiles of respiratory-virus detection rates

Figure 4 illustrates the temporal variation in respiratory-virus detection rates across the surveillance period. Each pathogen is represented by a separate time-series curve, with the vertical axis expressing mean reported detections per 1,000 specimens and the horizontal axis representing calendar time.

The figure enables comparisons of temporal circulation patterns among the respiratory-virus groups. Periods in which multiple pathogen curves increase simultaneously indicate potential temporal overlap in viral circulation. Conversely, differences in the timing of elevated detection rates may indicate that individual pathogens follow distinct seasonal trajectories.

The visualization is particularly useful for examining whether respiratory-virus activity is synchronized or temporally differentiated. For example, concurrent increases in several viral detection rates may indicate periods of intensified multi-pathogen circulation, whereas sequential peaks may suggest temporal displacement in the observed circulation patterns.

The temporal profiles should be interpreted in conjunction with the country-level surveillance coverage and specimen denominators. Changes in testing volume, reporting practices, and geographical contributions may influence the apparent magnitude of detection rates.

Scientific implication: Figure 4 provides a temporal epidemiological perspective on respiratory-virus co-circulation and supports the identification of seasonal overlap and potential temporal shifts in pathogen activity.

Shows the number of virus groups detected per country-week among observations with at least five pathogen groups reported. It summarizes the breadth of observed multi-virus circulation, not individual-level coinfection.

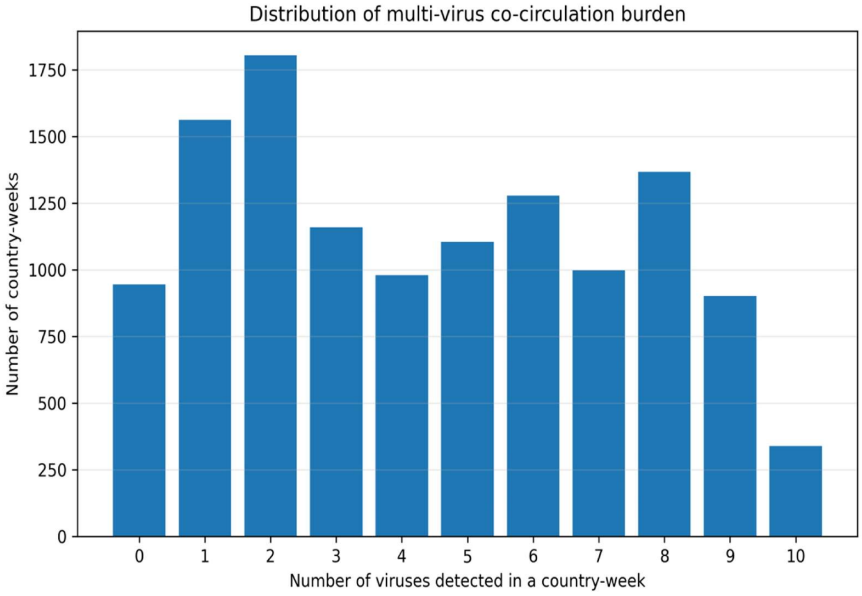

Figure 5. Distribution of multi-virus co-circulation burden

Figure 5 presents the distribution of the number of respiratory virus groups detected simultaneously within country-week observations for which at least five pathogen groups had available surveillance observations.

The horizontal axis represents the number of detected virus groups, while the vertical axis represents the corresponding number of country-week observations.

The distribution demonstrates that respiratory-virus circulation frequently involves multiple detected pathogens rather than being restricted to a single pathogen. In particular, country-weeks with two or more detected virus groups collectively account for 79.83% of the eligible observations. The largest individual category consists of country-weeks with two detected viruses (1,805 observations), followed by those with eight detected viruses (1,367 observations).

The occurrence of country-weeks with five or more detected virus groups further demonstrates the complexity of the observed respiratory-virus circulation landscape. Although observations involving all ten virus groups were less frequent, their presence indicates that extensive multi-pathogen detection occurred during certain surveillance periods.

These findings highlight the importance of examining respiratory-virus circulation using a multi pathogen framework rather than considering each virus independently. Nevertheless, the distribution represents country week level co-detection and should not be interpreted as the frequency of individual level coinfection.

Scientific implication: Figure 5 demonstrates the extent of multi-virus co-circulation and supports the need for integrated surveillance approaches that account for the simultaneous circulation of multiple respiratory pathogens.

7. Overall scientific interpretation of the findings

Taken together, the tables and figures demonstrate that respiratory-virus circulation is characterized by substantial pathogen-specific variation, frequent simultaneous detection, and heterogeneous relationships between viral detection patterns.

Three major findings emerge from the analysis.

First, influenza A and RSV exhibit particularly high detection frequencies, indicating their prominent representation in the surveillance dataset. However, substantial differences in pathogen-specific reporting coverage mean that these detection frequencies should not be interpreted as direct comparisons of population-level prevalence.

Second, the pairwise co-detection and correlation analyses reveal strong positive associations among several respiratory viruses. The particularly strong adenovirus–metapneumovirus association (OR = 20.33) and adenovirus parainfluenza correlation (ρ = 0.777) identify important pathogen combinations for further investigation.

Third, the co-circulation burden analysis demonstrates that multiple respiratory-virus groups are frequently detected within the same country week. This supports the importance of multi-pathogen surveillance and integrated epidemiological modelling.

Overall, these findings suggest that respiratory-virus circulation should be investigated through an integrated framework combining pathogen specific surveillance, pairwise association analysis, temporal modelling, and geographical comparisons.

Statistical interpretation: the reported analyses identify co-circulation patterns and associations. Claims of viral interference, facilitation, or mechanistic interaction require individual-level coinfection data or appropriately controlled longitudinal/experimental evidence.

9. Conclusion

9.1 Summary of principal findings

This study developed and applied an integrated, missingness-aware framework for characterising global respiratory virus co-circulation using 41,465 consolidated country week observations from 184 countries across a multi-year surveillance window spanning pandemic era and post-pandemic reporting conditions. Three principal findings emerge.

First, pathogen-specific circulation is markedly heterogeneous. Influenza A (78.49%) and RSV (63.51%) were the most frequently detected groups within their available reporting windows, whereas bocavirus (24.59%), other respiratory viruses (32.64%) and human coronavirus (39.14%) were least frequently detected. Cumulative detections and typical positive-observation intensity were dissociated: influenza A dominated in aggregate volume (4,395,375 detections), while rhinovirus exhibited the highest median count when positive (13). Because pathogen-specific denominators ranged from 9,341 to 36,391 country-weeks, these proportions describe reporting-window detection frequency rather than comparable population prevalence.

Second, pairwise analyses revealed a structured rather than uniform association landscape. The adenovirus–metapneumovirus pair showed the strongest binary co-detection association (OR = 20.33), while adenovirus–parainfluenza (ρ = 0.777), adenovirus metapneumovirus (ρ= 0.753) and metapneumovirus parainfluenza (ρ = 0.739) showed the strongest quantitative rank associations. Together with the rhinovirus links ($\rho \approx$ 0.68–0.70), these results delineate a coherent adenovirus metapneumovirus parainfluenza rhinovirus grouping, contrasted by weak associations involving bocavirus. The convergence of binary and count based measures

across the same pairs strengthens their descriptive credibility, while the consistently positive correlation structure indicates that shared circulation drivers plausibly contribute to the observed pattern Third, multi-pathogen detection was the norm rather than the exception: 79.83% of the 12,441 sufficiently reported country-weeks contained at least two detected virus groups, and 2.72% contained all ten. This supports treating respiratory-virus circulation as a multi-pathogen system and underscores the analytical incompleteness of single-pathogen surveillance summaries.

9.2 Methodological contribution

The study's principal methodological contribution is the explicit integration of data-quality assessment, pathogen-specific denominators, binary and quantitative association measures, multiplicity control, temporal profiling, burden quantification and network representation within a single reproducible architecture. Critically, missing pathogen observations were preserved as missing throughout, avoiding the artificial negative associations and underestimated circulation that zero-imputation would generate. Reporting complete-pair denominators, contingency-cell counts, FDR-adjusted q-values and precision estimates alongside effect sizes establishes an interpretive discipline that is often absent from pooled co-detection screens.

9.3 Interpretation boundaries

All conclusions are constrained by the ecological design. Country-week co-detection cannot establish that two viruses infected the same individual, nor that one pathogen facilitated or inhibited another. The pooled odds ratios and correlations are unadjusted and may partly reflect shared seasonality, geography, testing policy, multiplex panel composition, surveillance intensity and reporting completeness. The monthly indicator of detections per 1,000 specimens is a descriptive specimen-normalised measure, not a standardised positivity rate, unless numerator–denominator compatibility is verified. With 61.8% mean pathogen missingness and 7,805 low-testing country-weeks, apparent differences between pathogens, countries and periods may be surveillance artefacts. The proposed adjusted, lagged and network analyses (RO7–RO9) remain prospective and must not be reported as completed findings; with only ten nodes, network metrics are descriptive summaries rather than evidence of epidemiological importance.

9.4 Implications and future directions

Priority next steps are: (i) country and time adjusted mixed-effects or fixed-effects logistic models incorporating specimen volume, low testing status and reporting source, to determine which associations survive adjustment (H6); (ii) hemisphere and country stratified temporal and lagged analyses, since pooled monthly profiles may obscure opposing seasonal signals; (iii) sensitivity analyses across observation coverage thresholds, sentinel subsets and alternative association measures; and (iv) threshold-sensitivity testing of network centrality and community structure. Establishing biological interaction would require individual level coinfection, longitudinal or experimental evidence beyond the present dataset.

9.5 Concluding statement

Global respiratory virus surveillance reveals extensive, structured multi-pathogen co-detection, dominated by influenza A and RSV in detection frequency and by a tightly correlated adenovirus metapneumovirus parainfluenza rhinovirus grouping in association strength. The framework presented here characterises these patterns systematically while preserving the distinction between ecological statistical association and biological interaction, thereby providing a transparent and defensible foundation for integrated respiratory-virus monitoring and for prioritising pathogen combinations for subsequent epidemiological investigation.

References

[1] Jie, W., Wenyan, H., Chang, L., Jinghong, F., Wenyi, L., Shanshan, L., Ming, S. (2025). Epidemiological intricacies of respiratory pathogens: A single-center study on infection dynamics in Beijing, 2023–2024. *Frontiers in Public Health, 13*, 1581815. https://doi.org/10.3389/fpubh.2025.1581815.

[2] Petrova, V. N., Russell, C. A. (2018). The evolution of seasonal influenza viruses. *Nature Reviews Microbiology, 16*(1), 47–60. https://doi.org/10.1038/nrmicro.2017.118

[3] Baker, R. E., Mahmud, A. S., Wagner, C. E., Yang, W., Pitzer, V. E., Viboud, C., et al. (2019). Epidemic dynamics of respiratory syncytial virus in current and future climates. *Nature Communications, 10*, 5512. https://doi.org/10.1038/s41467-019-13562.

[4] Moriyama, M., Hugentobler, W. J., Iwasaki, A. (2020). Seasonality of respiratory viral infections. *Annual Review of Virology, 7*, 83–101. https://doi.org/10.1146/annurev-virology-012420-022445.

[5] Qian, T., Sprößig, W., Wang, J. (2011). Adaptive Fourier decomposition of functions in quaternionic Hardy spaces. *Mathematical Methods in the Applied Sciences, 35*(1), 43–64.

[6] Liu, H., Li, X., Sun, N., et al. (2026). Spatiotemporal instability of influenza seasonality during viral co-circulation. *npj Systems Biology and Applications, 12*, 103. https://doi.org/10.1038/s41540-026-00729-9.

[7] Zhao, Z., et al. (2023). Study on the prediction effect of a combined model of SARIMA and LSTM based on SSA for influenza in Shanxi Province, China. *BMC Infectious Diseases, 23*, 71.

[8] Krake, T., Klötzl, D., Hägele, D., Weiskopf, D. (2025). Uncertainty-aware seasonal-trend decomposition based on Loess. *IEEE Transactions on Visualization and Computer Graphics, 31*(3), 1496–1512.

[9] Qu, W., et al. (2025). A novel matching pursuit modeling strategy based on adaptive Fourier decomposition theory for predicting antigenic variation of influenza A (H1N1). *China CDC Weekly, 7*(14), 473–481.

[10] Cho, H. J., Rhee, J. E., Kang, D., Choi, E. H., Lee, N. J., Woo, S., Lee, J., Lee, S. W., Kim, E. J., Yun, K. W. (2024). Epidemiology of respiratory viruses in Korean children before and after the COVID-19 pandemic: A prospective study from National Surveillance System. *Journal of Korean Medical Science, 39*, e171.

[11] Mandelia, Y., Procop, G. W., Richter, S. S., Worley, S., Liu, W., Esper, F. (2021). Dynamics and predisposition of respiratory viral co-infections in children and adults. *Clinical Microbiology and Infection, 27*(4), 631.e1–631.e6.

[12] Huang, Q. S., Wood, T., Jelley, L., Jennings, T., Jefferies, S., Daniells, K., Nesdale, A., Dowell, T., Turner, N., Campbell-Stokes, P., et al. (2021). Impact of the COVID-19 nonpharmaceutical interventions on influenza and other respiratory viral infections in New Zealand. *Nature Communications, 12*, 1001.

[13] Yum, S., Hong, K., Sohn, S., Kim, J., Chun, B. C. (2021). Trends in viral respiratory infections during COVID-19 pandemic, South Korea. *Emerging Infectious Diseases, 27*(6), 1685–1688.

[14] Teo, K. W., Patel, D., Sisodia, S., Roland, D., Gaillard, E. A., & Tang, J. W. (2022). Rhinovirus persistence during the COVID-19 pandemic—Impact on pediatric acute wheezing presentations. *Journal of Medical Virology, 94*(11), 5547–5552.

[15] Kuitunen, I., Artama, M., Haapanen, M., Renko, M. (2021). Rhinovirus spread in children during the COVID-19 pandemic despite social restrictions A nationwide register study in Finland. *Journal of Medical Virology, 93*(10), 6063–6067.

[16] Dee, K., Goldfarb, D. M., Haney, J., Amat, J. A. R., Herder, V., Stewart, M., Szemiel, A. M., Baguelin, M., Murcia, P. R. (2021). Human rhinovirus infection blocks severe acute respiratory syndrome coronavirus 2 replication within the respiratory epithelium: Implications for COVID-19 epidemiology. *The Journal of Infectious Diseases, 224*(1), 31–38.

[17] Wu, A., Mihaylova, V. T., Landry, M. L., Foxman, E. F. (2020). Interference between rhinovirus and influenza A virus: A clinical data analysis and experimental infection study. *The Lancet Microbe, 1*(6), e254–e262.

[18] Iuliano, A. D., et al. (2018). Estimates of global seasonal influenza-associated respiratory mortality: A modelling

study. *The Lancet, 391*(10127), 1285–1300.

[19] Yu, H., et al. (2013). Characterization of regional influenza seasonality patterns in China and implications for vaccination strategies: Spatio-temporal modeling of surveillance data. *PLoS Medicine, 10*(11), e1001552.