

Metagenomic Sequencing Analysis of Respiratory Pathogens in the Jiangsu Region of China: A Multicenter Epidemiological Study (2025)

Yuan Mao

School of Biological Sciences and Medical Engineering, Southeast University

Zixuan Li

Clinical college of anhui medical university

Zhiqiang Li

Hefei KingMed Medical Laboratory

Ying Pan

School of Medical Technology, Anhui Institute of Medicine

Juan Gao

Hefei KingMed Medical Laboratory

Mingliang Zhang

Nanjing KingMed Medical Laboratory

Min Zhang

Nanjing KingMed Medical Laboratory

Jianfei Sun

sunzaghi@seu.edu.cn

School of Biological Sciences and Medical Engineering, Southeast University

Research Article

Keywords: Targeted metagenomic next-generation sequencing, Respiratory tract infection, Pediatric respiratory infection, Mixed infection, Epidemiology

Posted Date: June 9th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9728968/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at BMC Infectious Diseases on September 7th, 2026. See the published version at <https://doi.org/10.1186/s12879-026-14352-w>.

Abstract

Background

Acute respiratory infection (ARI) ranks as the leading cause of morbidity and mortality among children worldwide. Targeted metagenomic next-generation sequencing (tNGS) allows unbiased and comprehensive pathogen detection. This study aimed to investigate the epidemiological characteristics of respiratory pathogen infections in Jiangsu Province, China, using the tNGS platform.

Methods

A total of 28 736 clinical specimens were collected from patients with respiratory tract infections across 13 regions of Jiangsu Province between January and December 2025 and subjected to tNGS detection. All detected pathogens were classified into four categories: bacteria, viruses, atypical pathogens, and other pathogens. Stratified analyses were performed according to age, geographical region, season, and infection pattern (single infection versus mixed infection). The Chi-square test was applied for statistical analysis, with a two-tailed P value < 0.05 considered statistically significant.

Results

At least one respiratory pathogen was detected in 28,736 of 28,736 included specimens (100%); however, this reflects the high analytical sensitivity of tNGS and does not necessarily indicate clinically significant infection in every case. Bacteria exhibited the highest detection rate of 72.4%, followed by viruses (66.6%), atypical pathogens (6.3%), and other pathogens (3.2%). Mixed infections accounted for 68.7% of all cases, with the highest prevalence in children aged 1–4 years (73.5%) and the lowest in those aged ≥ 60 years (46.2%). The top five predominant pathogens were *Haemophilus influenzae* (26.4%), *Streptococcus pneumoniae* (23.0%), *Moraxella catarrhalis* (21.1%), *Staphylococcus aureus* (18.4%), and *Acinetobacter baumannii* (13.1%). Statistically significant differences were observed in pathogen distribution across age ($\chi^2 = 412.7$, $P < 0.001$), geographical region ($\chi^2 = 189.3$, $P < 0.001$), and season ($\chi^2 = 48.2$, $P < 0.001$). The most prevalent mixed infection combination was *Haemophilus influenzae* combined with *Streptococcus pneumoniae*.

Conclusion

This large-scale tNGS study reveals a complex pathogen spectrum of acute respiratory infections in East China, characterized by a high prevalence of mixed infections and significant disparities in age, regional, and seasonal distributions. The findings provide epidemiological evidence for precise clinical diagnosis and treatment as well as public health intervention strategies.

1. Introduction

1.1 Background

Acute respiratory infections (ARIs) represent the leading global burden of infectious diseases, responsible for approximately 4 million deaths annually, with children under 5 years of age being the most vulnerable population^[1]. In China, acute respiratory infections consistently rank among the top three causes of hospitalization in pediatric patients, imposing a substantial burden on both the healthcare system and families^[2–3].

Conventional pathogen detection methods, including microbial culture, antigen detection, and polymerase chain reaction (PCR), have inherent limitations. Microbial culture is time-consuming and exhibits low sensitivity for fastidious bacteria; antigen detection only targets specific pathogens, frequently leading to missed detection of unknown pathogens. Although PCR assays possess high sensitivity, they require pre-designed target primers and are prone to selection bias^[4–6]. These technical deficiencies ultimately result in a considerable number of missed or delayed diagnoses of respiratory tract infections.

1.2 Metagenomic Sequencing in Clinical Practice

Targeted next-generation sequencing (tNGS) combines ultra-multiplex PCR enrichment with high-throughput sequencing, providing a practical compromise that improves detection sensitivity while reducing assay costs^[7–9]. tNGS enables the simultaneous identification of bacteria, viruses, fungi, parasites and other pathogens directly from clinical specimens without microbial cultivation, conferring unique advantages in the diagnosis of immunocompromised patients, critically ill children, and outbreaks of unexplained infections. Accumulating large-sample tNGS data provide an important complement to traditional surveillance approaches for population-level epidemiological research.

1.3 Study Objectives

Although the clinical application of targeted metagenomic next-generation sequencing (tNGS) has become increasingly widespread, large-scale epidemiological data on respiratory pathogens in East China remain scarce. The present study aimed to: (1) clarify the composition and detection rates of respiratory pathogens across 13 regions of Jiangsu Province; (2) analyze the distribution patterns of pathogens in terms of age, geographical region, and season; (3) evaluate the incidence and clinical characteristics of co-infections based on 2025 tNGS data.

2. Materials and Methods

2.1 Study Design and Population

This study was designed as a retrospective, multicenter, cross-sectional analysis based on clinical targeted metagenomic next-generation sequencing (tNGS) data. A total of 28 736 respiratory specimens

from patients with acute respiratory infections, collected between January and December 2025, were enrolled. Specimens were sourced from 13 regions across Jiangsu Province, including Suzhou, Nanjing, Yangzhou, Lianyungang, Wuxi, Changzhou, Taizhou, Nantong, Xuzhou, Huai'an, Suqian, Yancheng, and other collaborative medical centers. Inclusion criteria: patients aged 0–18 years with clinically diagnosed acute respiratory infection. Exclusion criteria: unqualified specimens, repeated submissions, and patients with severe chronic or immunocompromising conditions. All specimens were submitted for clinical respiratory pathogen tNGS testing as part of routine diagnostic workup.

2.2 Specimen Collection and Processing

Respiratory specimens (including nasopharyngeal swabs ~ 85%, sputum ~ 10%, bronchoalveolar lavage fluid/throat swabs ~ 5%) were collected according to standard clinical protocols and stored at -80 degrees C until processing. DNA and RNA were co-extracted using a commercial kit according to the manufacturer's instructions. Library construction and sequencing were performed on an Illumina platform with a minimum depth of 20 million reads per sample.

2.3 Pathogen Detection and Classification

Sequencing data were processed using a validated bioinformatics pipeline. Low-quality reads ($Q < 30$) and human host reads were removed by alignment to the human reference genome (GRCh38). Remaining reads were classified against a comprehensive respiratory pathogen reference database. A pathogen was considered positive when normalized coverage $\geq 0.1\%$ and effective reads ≥ 50 . Pathogens were classified into four categories: (1) Bacteria, (2) Virus, (3) Atypical Pathogen (Mycoplasma, Chlamydia, Legionella), and (4) Others. Non-pathogenic detections (e.g., antibiotic resistance genes, 'no pathogen detected' results) were excluded.

2.4 Data Extraction and Stratification

Patients were stratified into seven age groups: < 1 year, 1–4 years, 5–13 years, 14–17 years, 18–29 years, 30–59 years, and ≥ 60 years, reflecting infancy, early childhood, school age, adolescence, and adulthood. Geographic region was assigned based on hospital location. Temporal stratification included four seasons (Spring: March–May; Summer: June–August; Autumn: September–November; Winter: December–February) and four quarters. Infection pattern was classified as 'single infection' (one pathogen category detected) or 'mixed infection' (≥ 2 pathogen categories detected in a single specimen).

2.5 Statistical Analysis

Descriptive statistics summarized pathogen detection rates across strata. Detection rate was defined as the number of specimens positive for a given pathogen divided by the total number of specimens ($N = 28,736$), expressed as a percentage. For mixed-infection analysis, co-infection pairs were counted when both pathogens appeared in the same specimen. Statistical significance of differences across strata was assessed using the chi-square (chi-squared) test of independence on contingency tables.

Significance threshold was $\alpha = 0.05$. Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All analyses were performed using Python 3.9+.

2.6 Ethical Considerations

This study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of Guangzhou KingMed Diagnostics Center (Approval No.: 2026083; Date: May 3, 2026). The study used retrospectively collected, de-identified clinical tNGS data; the requirement for additional written informed consent specifically for this research analysis was waived by the IRB due to the retrospective and anonymized nature of the data (Expedited Review). Participants under 16 years of age constituted the majority of this study population (at least $N = 24,379$, 84.8% were aged < 14 years, with an estimated 87–92% aged < 16 years). For these pediatric participants, informed consent for clinical respiratory pathogen testing (including tNGS) had been obtained from their parents or legal guardians at the time of specimen collection as part of routine diagnostic procedures, in accordance with institutional policies and applicable national regulations. The standardized test requisition form included a dedicated section for parental or legal guardian consent. We acknowledge that while consent forms were designed to capture this information, retrospective review revealed that a proportion of forms had incomplete entries in the consent section. This limitation reflects real-world challenges in the documentation of routine clinical consent processes and is further addressed in the Limitations section.

3. Results

3.1 Overall Pathogen Detection

A total of 100% of the 28 736 specimens tested by tNGS yielded at least one respiratory pathogen. It should be noted that tNGS detects microbial nucleic acids at high sensitivity, and these findings may include colonizing organisms that are not necessarily the causative pathogens of the acute respiratory illness. In terms of pathogen composition, bacteria were the most prevalent, detected in 20 796 cases (72.4%), followed by viruses in 19 150 cases (66.6%), atypical pathogens in 1 820 cases (6.3%), and other pathogens in 906 cases (3.2%). The summed detection rates of all pathogen categories exceeded 100%, which was consistent with a high prevalence of mixed infections. Mixed infections accounted for 68.7% of all cases, with the highest prevalence in children aged 1–4 years (73.5%) and the lowest in those aged ≥ 60 years (46.2%).

Single infection accounted for 31.3% (8993/28736) of all cases, while mixed infections constituted 68.7% (19743/28736) (Fig. 1).

3.2 Top 20 Pathogens

The 20 most frequently detected pathogens are summarized in Table 1 and Fig. 2. Among the top five pathogens, four were bacteria: *Haemophilus influenzae* (7,594; 26.4%, a major cause of pediatric pneumonia), *Streptococcus pneumoniae* (6,615; 23.0%, the leading pathogen of community-acquired pneumonia), *Moraxella catarrhalis* (6,072; 21.1%, common in young children), and *Staphylococcus*

aureus (5,294; 18.4%, associated with severe infections). Acinetobacter baumannii ranked fifth overall (3,751; 13.1%, linked to hospital-acquired infections). Among viruses, Rhinovirus A was the most prevalent (3,188; 11.1%), followed by RSV-A (2,825; 9.8%), Rhinovirus C (2,296; 8.0%), and Human Metapneumovirus (1,821; 6.3%). Atypical pathogens were less common, with Chlamydia pneumoniae (915; 3.2%) being the most frequent.

Table 1

Top 20 Respiratory Pathogens Detected by Metagenomic Sequencing in 28,736 Pediatric Specimens.

Rank	Pathogen	Detections (n)	Detection Rate (%)	Category
1	Haemophilus influenzae	7,594	26.43	Bacteria
2	Streptococcus pneumoniae	6,615	23.02	Bacteria
3	Moraxella catarrhalis	6,072	21.13	Bacteria
4	Staphylococcus aureus	5,294	18.42	Bacteria
5	Acinetobacter baumannii	3,751	13.05	Bacteria
6	Rhinovirus A	3,188	11.09	Virus
7	Respiratory Syncytial Virus A	2,825	9.83	Virus
8	Rhinovirus C	2,296	7.99	Virus
9	Streptococcus anginosus group	2,059	7.17	Bacteria
10	Human Metapneumovirus	1,821	6.34	Virus
11	Cytomegalovirus (CMV)	1,708	5.94	Virus
12	Respiratory Syncytial Virus B	1,647	5.73	Virus
13	Klebsiella pneumoniae	1,282	4.46	Bacteria
14	Influenza A H1N1 (2009)	1,274	4.43	Virus
15	Streptococcus intermedius	1,230	4.28	Bacteria
16	Human Parainfluenza Virus 3	1,063	3.70	Virus
17	Herpes Simplex Virus 1 (HSV-1)	954	3.32	Virus
18	Chlamydia pneumoniae	915	3.18	Atypical
19	Epstein-Barr Virus (EBV)	881	3.07	Virus
20	Streptococcus pyogenes	872	3.03	Bacteria
Data are presented as n (%). Detection rate = number of detections / total specimens (N = 28,736) x 100%				

3.3 Age-Stratified Analysis

Significant age-related variations were observed (chi-squared test: $p < 0.001$). The 1–4 year age group contributed the largest proportion ($n = 12,098$; 42.1%), followed by 5–13 years ($n = 8,121$; 28.3%) and < 1 year ($n = 4,160$; 14.5%). Patients aged 18–29 years were the smallest subgroup ($n = 243$; 0.8%), reflecting the pediatric focus of clinical mNGS testing (Table 2, Fig. 3). RSV-A predominated in infants < 1 year ($> 40\%$), while *H. influenzae* and *S. pneumoniae* dominated in children aged 1–4 years. Rhinovirus was prevalent across ages 1–13 years, associated with wheezing and asthma exacerbations.

Table 2
Specimen Distribution by Age Group and Infection Pattern.

Age Group	Specimens (n)	Proportion (%)	Infection Pattern
< 1 yr	4,160	14.5	Mixed: 3,012 (72.4%); Single: 1,148 (27.6%)
1–4 yr	12,098	42.1	Mixed: 8,891 (73.5%); Single: 3,207 (26.5%)
5–13 yr	8,121	28.3	Mixed: 5,592 (68.9%); Single: 2,529 (31.1%)
14–17 yr	394	1.4	Mixed: 247 (62.7%); Single: 147 (37.3%)
18–29 yr	243	0.8	Mixed: 142 (58.4%); Single: 101 (41.6%)
30–59 yr	1,103	3.8	Mixed: 651 (59.0%); Single: 452 (41.0%)
≥ 60 yr	2,617	9.1	Mixed: 1,208 (46.2%); Single: 1,409 (53.8%)

Mixed infection: ≥ 2 pathogen categories detected in a single specimen; Single infection: 1 pathogen category detected. Chi-squared test: $p < 0.001$.

3.4 Seasonal Distribution

Specimen distribution across seasons: Spring (Q1: $n = 9,276$; 32.3%), Summer (Q2: $n = 6,996$; 24.3%), Autumn (Q3: $n = 6,433$; 22.4%), Winter (Q4: $n = 6,031$; 21.0%). Seasonal variation was statistically significant (chi-squared = 48.2, $p < 0.001$). Bacterial pathogens predominated across all seasons (72.3–72.9%). Viral detection showed modest seasonal fluctuation: Rhinovirus peaked in spring (March–May) and autumn (September–November), while RSV-A showed a winter-predominant pattern (November–February), with the highest rate in December, consistent with its known epidemiology (Fig. 4).

3.5 Regional Distribution

Suzhou contributed the largest number of specimens ($n = 8,133$; 28.3%), followed by Nanjing ($n = 4,174$; 14.5%), Yangzhou ($n = 3,578$; 12.4%), and Lianyungang ($n = 2,484$; 8.6%). Significant regional heterogeneity was observed (chi-squared = 189.3, $p < 0.001$). Suzhou showed the highest bacterial detection rate (79.3%), above the regional average (72.4%), possibly related to dense healthcare resources, high population mobility, and humid climate in southern Jiangsu (Table 3, Fig. 5).

Table 3
Geographic Distribution of Specimens and Infection Patterns.

Region	Specimens (n)	Proportion (%)	Infection Pattern Distribution
Suzhou	8,133	28.3	Mixed: 6,102 (75.0%); Single: 2,031 (25.0%)
Nanjing	4,174	14.5	Mixed: 2,934 (70.3%); Single: 1,240 (29.7%)
Yangzhou	3,578	12.4	Mixed: 2,451 (68.5%); Single: 1,127 (31.5%)
Others*	3,249	11.3	Mixed: 2,189 (67.4%); Single: 1,060 (32.6%)
Lianyungang	2,484	8.6	Mixed: 1,679 (67.6%); Single: 805 (32.4%)
Wuxi	1,957	6.8	Mixed: 1,318 (67.3%); Single: 639 (32.7%)
Changzhou	1,879	6.5	Mixed: 1,252 (66.6%); Single: 627 (33.4%)
Taizhou	1,239	4.3	Mixed: 805 (65.0%); Single: 434 (35.0%)
Nantong	1,051	3.7	Mixed: 692 (65.8%); Single: 359 (34.2%)
Xuzhou	607	2.1	Mixed: 374 (61.6%); Single: 233 (38.4%)
Huai'an	287	1.0	Mixed: 196 (68.3%); Single: 91 (31.7%)
Suqian	73	0.3	Mixed: 45 (61.6%); Single: 28 (38.4%)
Yancheng	25	0.1	Mixed: 16 (64.0%); Single: 9 (36.0%)

* Others: aggregate of regions with < 50 specimens per site. Chi-squared test: $p < 0.001$.

3.6 Single vs. Mixed Infection Patterns

Of the 28,736 specimens, 8,993 (31.3%) yielded a single pathogen category (single infection) and 19,743 (68.7%) yielded multiple categories (mixed infection). The prevalence of mixed infections was highest in the 1–4 year group (73.5%) and lowest in the ≥ 60 year group (46.2%), suggesting that young children are more likely to harbor multiple simultaneous pathogen categories (Fig. 6). It should be noted that “mixed infection” in this study was defined at the pathogen category level (i.e., ≥ 2 of the following four categories: bacteria, viruses, atypical pathogens, or others), which may differ from species-level co-infection definitions used in other studies. This category-based approach may result in a higher mixed infection rate compared to studies defining co-infection at the species level.

3.7 Co-infection Pair Analysis

The ten most frequent co-infection pairs are presented in Table 4 and Fig. 7. The most common pair was *Haemophilus influenzae* + *Streptococcus pneumoniae* ($n = 215$; common nasopharyngeal co-colonizers, pathogenic during immune compromise). Viral-bacterial pairs were also prominent: RSV-A + *S. pneumoniae* ($n = 205$; viral priming followed by secondary bacterial infection, higher severity risk) and

RSV-A + M. catarrhalis (n = 176; common in young children), consistent with viral infection predisposing to secondary bacterial colonization.

Table 4
Top 10 Most Frequent Co-infection Pathogen Pairs.

Rank	Co-infection Pair	Count (n)	Relative Note
1	Haemophilus influenzae + Streptococcus pneumoniae	215	Most frequent bacterial co-infection pair
2	RSV-A + Streptococcus pneumoniae	205	Viral-bacterial co-infection
3	RSV-A + Moraxella catarrhalis	176	Viral-bacterial co-infection
4	Haemophilus influenzae + Rhinovirus A	167	Viral-bacterial co-infection
5	Moraxella catarrhalis + Streptococcus pneumoniae	154	Bacterial co-infection
6	RSV-A + Haemophilus influenzae	131	Viral-bacterial co-infection
7	Moraxella catarrhalis + Rhinovirus C	129	Viral-bacterial co-infection
8	Haemophilus influenzae + Acinetobacter baumannii	120	Bacterial co-infection
9	Haemophilus influenzae + Rhinovirus C	118	Viral-bacterial co-infection
10	Klebsiella pneumoniae + Acinetobacter baumannii	116	Bacterial co-infection

Co-infection pairs were identified from all mixed-infection records (n = 19,743). Only pairs with > = 2 distinct pathogens are included.

3.8 Monthly Heatmap of Pathogen Dynamics

A comprehensive monthly heatmap (Fig. 8) revealed dynamic shifts in pathogen profiles. Rhinovirus A and Rhinovirus C showed pronounced Spring (March–May) and Autumn (September–November) peaks. RSV-A and RSV-B exhibited a clear Winter predominance (November–February), aligning with their known epidemiology. Haemophilus influenzae and Streptococcus pneumoniae showed relatively stable detection rates throughout the year.

4. Discussion

4.1 Pathogen Landscape and Clinical Implications

This study performed a large-scale analysis of 28 736 pediatric respiratory specimens in East China based on targeted metagenomic next-generation sequencing (tNGS), and comprehensively characterized

the epidemiological profiles of respiratory pathogens in 2025. Metagenomic sequencing (including mNGS and tNGS) for respiratory specimens exhibits a significantly higher detection yield compared with microbial culture, PCR, and serological assays, and bronchoalveolar lavage fluid (BALF) specimens demonstrate far greater clinical value than sputum samples. Targeted tNGS features higher sensitivity and lower cost, making it suitable for the screening of common respiratory pathogens^[10]. The 100% pathogen detection rate indicates a high analytical sensitivity of the tNGS platform; however, the clinical significance of individual pathogen detections requires interpretation in conjunction with clinical presentation in symptomatic children. The mixed infection rate reached as high as 68.7%, which challenges the traditional clinical concept of single-pathogen pathogenesis and highlights the necessity of comprehensive pathogen detection^[11]. The top five predominant pathogens were consistent with the epidemiological characteristics of nasopharyngeal bacterial colonization in children^[12–14].

4.2 Age as a Determinant of Pathogen Profile

The age distribution of the study population showed statistically significant differences, with children aged 1–4 years accounting for more than 42% of all enrolled individuals. This distribution pattern is highly consistent with the global age-related disease burden of severe acute respiratory infections. From a clinical perspective, children aged 1–4 years are in a critical period of rapid but immature immune development, accompanied by weakened respiratory mucosal barrier function and relatively low resistance to various respiratory pathogens. Meanwhile, children in this age group have an expanded scope of daily activities, which increases the probability of pathogen exposure, thereby rendering them a high-risk population for severe acute respiratory infections^[15]. These findings further confirm the age-dependent pattern of pediatric respiratory infections and provide an important age-targeted basis for precise prevention and stratified clinical intervention of severe acute respiratory infections. It is recommended to prioritize the monitoring and early intervention strategies for children aged 1–4 years, so as to reduce the risk of disease progression to severe cases.

In terms of the age distribution of viral detection, respiratory syncytial virus A (RSV-A) exhibited the highest detection rate (> 40%) in infants aged less than 1 year, which is fully consistent with the disease burden characteristics of severe bronchiolitis in this age group^[16]. Neonates and infants possess unique respiratory anatomical features, including narrow airways and abundant mucosal blood vessels. Once infected with RSV-A, airway mucosal congestion and edema may occur, potentially resulting in airway obstruction and subsequent progression to severe bronchiolitis, which may be life-threatening in severe cases^[17]. This finding indicates that RSV-A is one of the major pathogens responsible for severe respiratory infections in infants under 1 year of age. Clinically, RSV-A infection should be prioritized in infants presenting with acute respiratory symptoms, with timely pathogen detection and targeted treatment initiated accordingly. Meanwhile, preventive interventions against RSV (such as vaccination and passive immunization) should be explored for high-risk infants to alleviate the corresponding disease burden.

Different from the age-concentrated distribution of RSV-A, rhinovirus infection showed a broader age distribution, with a prominent peak among children aged 1–13 years, and was closely associated with

wheezing episodes and acute asthma exacerbations^[18]. On the one hand, children aged 1–13 years span from early childhood to adolescence, with increasing opportunities for exposure to external environments. As a prevalent respiratory pathogen, rhinovirus is easily transmitted via respiratory droplets and close contact, leading to an elevated infection rate^[19]. On the other hand, rhinovirus infection has been associated with inflammatory responses in the airway mucosa, damage airway epithelial cells, and induce airway hyperresponsiveness. For children with underlying airway hypersensitivity or asthma, rhinovirus infection may precipitate wheezing attacks or acute asthma exacerbations, which is in line with previous findings regarding the association between rhinovirus infection and pediatric wheezing as well as acute asthma episodes^[20].

4.3 Seasonal Characteristics and Epidemiological Surveillance

The seasonal distribution of viral infections in this study was basically consistent with previously reported data^[21–22], which further verified the reliability of the research findings. Rhinovirus infection exhibited a peak incidence in spring and autumn, which is in line with the epidemiological characteristics of rhinovirus in temperate regions—year-round prevalence with concentrated outbreaks in spring and autumn^[23]. Such seasonal distribution differences may be, speculatively, closely associated with drastic temperature fluctuations in spring and autumn, reduced adaptability of children’s respiratory mucosa, and increased opportunities for pathogen transmission caused by heightened crowd aggregation activities.

In contrast, respiratory syncytial virus A (RSV-A) showed a typical winter peak distribution. This feature not only conforms to the established seasonal epidemic pattern of RSV infection^[24] but also validates the internal consistency of viral detection data obtained via tNGS technology in this study.

The seasonal distribution characteristics of viral infections provide an important time-targeted basis for the seasonal prevention and control of pediatric respiratory tract infections. The findings of this study can further improve the regional surveillance system for pediatric respiratory viral infections, supplement the seasonal epidemiological data of different viruses, and offer scientific support for the formulation of clinical seasonal intervention strategies. For instance, targeted prevention and monitoring measures for pediatric respiratory tract infections can be strengthened in spring and autumn to reduce clustered transmission in response to the peak prevalence of rhinovirus during these seasons. Regarding the winter predominance of RSV-A, it is feasible to optimize RSV screening and intervention protocols for infants in winter, rationally arrange the vaccination schedule, and enhance the pertinence and effectiveness of preventive interventions. Ultimately, this can further reduce the disease burden of pediatric respiratory tract infections across different seasons.

4.4 Regional Characteristics and Epidemiological Surveillance

The results of this study revealed marked regional differences in the volume of submitted specimens and bacterial detection rates across prefecture-level cities in Jiangsu Province. Suzhou had the largest number of submitted specimens in the province, with a bacterial detection rate of 79.3%, which was higher than the provincial average of 72.4%. Significant statistical heterogeneity was observed in the pathogen detection distribution among different regions ($\chi^2=189.3$, $P < 0.001$), indicating that the regional differentiation of bacterial epidemic characteristics was genuine rather than attributable to sampling error.

The relatively high bacterial detection rate in Suzhou can be explained by multiple factors. First, Suzhou enjoys a well-developed layout of medical resources, with dense coverage of tertiary hospitals, maternal and child health institutions, and third-party laboratory testing platforms. The popularization of grassroots pathogen screening is high, and the proportion of residents taking the initiative for medical consultation and health screening is greater. This enrolls more patients with mild and subclinical infections, which may contribute to a higher observed positive detection rate. Second, Suzhou has a large resident population and a highly concentrated floating population. Frequent population mobility and dense crowd aggregation facilitate community transmission and cryptic circulation of pathogenic bacteria. Third, the southern Jiangsu region features a humid climate and a prolonged plum rain season. The warm and humid environment is conducive to the growth and colonization of pathogenic bacteria.

The regional heterogeneity of bacterial detection rates across Jiangsu Province results from the combined effects of medical resource allocation, population mobility patterns, specimen composition, and geographical environmental conditions. The findings of this study can provide an epidemiological reference for implementing localized pathogen surveillance, primary screening, and formulating targeted infection prevention and control strategies in Jiangsu Province.

4.5 Clinical Significance of Mixed Infections

The overall mixed infection rate in this study reached 68.7%, which was markedly higher than the range of 14%–30% reported in previous studies adopting conventional culture methods^[25–26]. This suggests that the higher mixed infection rate may partially reflect the superior sensitivity of tNGS in pathogen detection. Limited by culture conditions and intrinsic growth properties of microorganisms, conventional culture approaches frequently fail to detect fastidious or low-abundance pathogens. In contrast, tNGS enables simultaneous screening for a broad spectrum of pathogens and substantially improves the detection efficiency of mixed infections.

Among bacterium-bacterium coinfections, the combination of *Haemophilus influenzae* plus *Streptococcus pneumoniae* ranked among the most prevalent^[27]. Multiple lines of evidence indicate that *Streptococcus pneumoniae*, *Haemophilus influenzae*, and *Moraxella catarrhalis* constitute a common bacterial commensal consortium in the pediatric respiratory tract^[28]. In addition, virus-bacterium coinfection patterns are prevalent in respiratory pathogen infections^[29–30]. Respiratory viral infection can impair the respiratory mucosal barrier and disrupt local immune homeostasis, creating

favorable conditions for bacterial colonization and invasion. This observation highlights the need for clinical vigilance against viral–bacterial co-infections in young children with acute respiratory infections.

Collectively, the age-specific distribution, seasonal patterns, and polymicrobial infection characteristics observed in this study reflect the intrinsic transmission and pathogenic profiles of respiratory pathogens. These findings enrich the epidemiological database of pediatric respiratory infections and provide valuable evidence for precise clinical prevention, etiological diagnosis, and individualized therapeutic strategies.

4.6 Strengths and Limitations

Strengths include the large sample size (N = 28,736), comprehensive tNGS-based detection covering a broad spectrum of respiratory pathogens, multi-center geographic coverage across 13 regions within Jiangsu Province, full-year temporal sampling encompassing all four seasons, and a standardized bioinformatics and statistical analysis pipeline.

Limitations of this study should be carefully considered:

First, the retrospective design introduces inherent selection bias. As this study utilized clinically referred specimens, the cohort is subject to referral bias - tNGS testing was primarily ordered for patients with clinical suspicion of atypical or severe infection, failure of conventional diagnostic methods, or complex medical histories. This likely over-represents diagnostically challenging or unusual cases and may not fully reflect the true pathogen distribution in community-acquired ARIs. Additionally, the uneven geographic distribution of specimens (Suzhou alone contributing 28.3% of all samples, while some regions contributing < 2%) further introduces regional selection bias, potentially over-representing pathogen epidemiology in southern Jiangsu.

Second, tNGS cannot reliably differentiate between active infection and asymptomatic colonization. The platform detects microbial nucleic acid sequences above predefined thresholds (normalized coverage $\geq 0.1\%$; effective reads ≥ 50) regardless of whether the organism represents the causative pathogen, a colonizing commensal, an environmental contaminant, or a transiently present organism. This is particularly relevant for bacteria such as *H. influenzae* (26.4%), *S. pneumoniae* (23.0%), and *M. catarrhalis* (21.1%), which are well-documented nasopharyngeal commensals in children, especially in the 1–4 year age group where respiratory carriage rates are known to be highest. The high overall bacterial detection rate (72.4%) and mixed infection rate (68.7%) may therefore partly reflect the natural respiratory microbiota of the study population rather than solely representing pathogenic involvement in each acute illness episode.

Third, detailed clinical information was limited. Available data were restricted to basic demographic variables (age, sex, geographic region, collection date) and specimen type. Critical clinical parameters were not available, including: disease severity indicators (e.g., need for supplemental oxygen, intensive care unit admission, mechanical ventilation, length of hospital stay); antimicrobial treatment history prior to and following tNGS testing; vaccination status (particularly against *S. pneumoniae*, *H. influenzae* type

b, and RSV); presence of underlying comorbidities (e.g., asthma, congenital heart disease, immunodeficiency); and definitive clinical outcomes (e.g., recovery, complications, mortality). This limits our ability to correlate specific pathogen detection profiles with clinical severity, treatment response, or prognosis.

Fourth, the absence of a healthy control group prevents determination of baseline carriage rates. Without asymptomatic controls matched by age, region, and season, we cannot determine whether detected pathogens represent true infection or background respiratory flora. Inclusion of a healthy control cohort would enable calculation of odds ratios for individual pathogens and help differentiate pathogenic organisms from benign commensals.

Fifth, data on antimicrobial exposure prior to specimen collection were not available. Prior antibiotic use may suppress growth and reduce detectability of susceptible bacterial pathogens (potentially leading to false-negative results for organisms such as *S. pneumoniae*) while simultaneously selecting for resistant strains (potentially over-representing organisms such as *A. baumannii*, 13.1%). The direction and magnitude of this bias cannot be quantified in the current dataset.

Sixth, the tNGS technology has inherent technical limitations. The targeted panel approach employed in this study covers a predefined set of respiratory pathogens; emerging, novel, or unexpected pathogens outside the reference database would not be detected. Additionally, semi-quantitative interpretation based on normalized coverage and read counts has not been fully harmonized across different tNGS platforms and bioinformatics pipelines, and inter-laboratory variability in wet-lab protocols (e.g., nucleic acid extraction efficiency, library preparation methods) may affect detection thresholds and limit cross-study comparability.

Seventh, generalizability is constrained by geography and population. Although specimens were collected from 13 regions across Jiangsu Province, the findings may not be directly generalizable to other provinces in China with different climates, population densities, socioeconomic characteristics, or pathogen epidemiology. Furthermore, international generalizability is likely limited given regional variations in vaccination policies, antimicrobial stewardship practices, and healthcare access patterns. Replication studies in other geographic regions are warranted.

4.7 Future Directions

Future studies should integrate prospective, multi-center clinical data with tNGS profiles to build predictive models for severe ARI outcomes. Longitudinal cohort studies with healthy controls will help establish causality and differentiate colonization from infection. The integration of metagenomics with host transcriptomic and proteomic profiling holds promise for developing multi-modal signatures of disease severity. Application of mNGS to outbreak surveillance in pediatric settings may enable earlier detection of emerging respiratory pathogens and antimicrobial-resistant strains.

5. Conclusions

Several important caveats should be considered when interpreting these findings. First, the retrospective design using clinically referred specimens introduces potential selection bias, as tNGS testing was typically ordered for patients with suspected atypical or severe infection, or those who failed to respond to empirical treatment. Second, tNGS detects nucleic acid sequences regardless of whether the organism represents active infection, asymptomatic colonization, or environmental contamination - this is particularly relevant for *H. influenzae*, *S. pneumoniae*, and *M. catarrhalis*, which are known nasopharyngeal commensals in young children. Third, the absence of healthy control groups and clinical outcome data precludes causal inference between detected pathogens and disease severity. Fourth, the highly skewed geographic distribution (Suzhou contributing 28.3% of all specimens) means that regional patterns, particularly for cities with smaller sample sizes, should be interpreted cautiously. Despite these limitations, this study provides the most comprehensive tNGS-based characterization to date of the respiratory pathogen landscape in Jiangsu Province and offers a foundation for future prospective validation studies.

This multi-center, cross-sectional tNGS study of 28 736 pediatric respiratory specimens from Jiangsu provinces provides the most comprehensive characterization to date of the respiratory pathogen landscape in this population. Key conclusions are: (1) all specimens yielded at least one pathogen, with bacteria (72.4%) and viruses (66.6%) being predominant; (2) mixed infections were highly prevalent (68.7%), emphasizing the polymicrobial nature of pediatric ARIs; (3) distinct age-related, regional, and seasonal patterns were identified ($p < 0.001$); (4) *Haemophilus influenzae*, *Streptococcus pneumoniae*, and *Moraxella catarrhalis* were the most frequently detected pathogens; (5) co-infection pairs between *H. influenzae* and *S. pneumoniae*, and between RSV-A and bacterial pathogens, were the most common. These findings underscore the value of comprehensive, hypothesis-free diagnostic approaches such as tNGS in understanding and managing pediatric respiratory infections in eastern China.

Declarations

Funding

This work was supported by the 2024 Anhui Provincial Higher Education Quality Engineering Project (Grant No. 2024xdcy036).

Author Contribution

Yuan Mao: Data curation, Formal analysis, Writing – original draft, Funding acquisition; Zixuan Li: Data curation, Investigation; Zhiqiang Li: Data curation, Methodology; Juan Gao: Software, Validation; Ying Pan: Resources, Supervision; Mingliang Zhang: Project administration; Min Zhang: Visualization; Jianfei Sun: Conceptualization, Writing – review & editing.

References

1. JOSÉ R J. Respiratory infections: a global burden[J]. *Annals Res Hosp*. 2018;2:12–12.
2. GRAHAM F. Daily briefing: What's causing a pneumonia surge in China[J/OL]. *Nature*. 2023. <http://dx.doi.org/10.1038/d41586-023-03761-5>.
3. WEN S, LIN Z, ZHANG Y, et al. The Epidemiology, Molecular, and Clinical of Human Adenoviruses in Children Hospitalized With Acute Respiratory Infections[J]. *Front Microbiol*. 2021;12. <http://dx.doi.org/10.3389/fmicb.2021.629971>.
4. AMJAD M. An Overview of the Molecular Methods in the Diagnosis of Gastrointestinal Infectious Diseases[J]. *International Journal of Microbiology*, 2020, 2020: 1–13.
5. DUAN H, LI X. The diagnostic value of metagenomic next-generation sequencing in infectious diseases[J]. *BMC Infect Dis*. 2021;21(1). <http://dx.doi.org/10.1186/s12879-020-05746-5>.
6. ZHENG Y, QIU X, WANG T, et al. The Diagnostic Value of Metagenomic Next-Generation Sequencing in Lower Respiratory Tract Infection[J]. *Front Cell Infect Microbiol*. 2021;11. <http://dx.doi.org/10.3389/fcimb.2021.694756>.
7. Tao Y, Zhou Z, Duan Y, et al. Diagnostic Value of Targeted Next-generation Sequencing in Pulmonary Mycobacterial Infections. *CURR MED SCI*. 2024;44:947–53. <https://doi.org/10.1007/s11596-024-2937-4>.
8. Chen Y, Fan L, Ren Z, et al. Sensitive diagnosis of paucibacillary tuberculosis with targeted next-generation sequencing: a molecular diagnostic study. *BMC Med*. 2025;23:178. <https://doi.org/10.1186/s12916-025-03996-1>.
9. Li C, Zhang X, Liu P, et al. Targeted next-generation sequencing characterization of respiratory pathogens in children with acute respiratory infection. *BMC Infect Dis*. 2025;25:675. <https://doi.org/10.1007/s12879-025-11051-w>.
10. Yi L, Tan L, Long Q, et al. Comparative diagnostic performance of metagenomic and two targeted sequencing methods in lower respiratory infection. *Sci Rep*. 2025;15:27365. <https://doi.org/10.1038/s41598-025-11834-w>.
11. Jain S, Self WH, Wunderink RG, Fakhran S, Balk R, Bramley AM, Reed C, Grijalva CG, Anderson EJ, Courtney DM, Chappell JD, Qi C, Hart EM, Carroll F, Trabue C, Donnelly HK, Williams DJ, Zhu Y, Arnold SR, Ampofo K, Waterer GW, Levine M, Lindstrom S, Winchell JM, Katz JM, Erdman D, Schneider E, Hicks LA, McCullers JA, Pavia AT, Edwards KM, Finelli L. CDC EPIC Study Team. Community-Acquired Pneumonia Requiring Hospitalization among U.S. Adults. *N Engl J Med*. 2015;373(5):415 – 27. doi: 10.1056/NEJMoa1500245. Epub 2015 Jul 14. PMID: 26172429; PMCID: PMC4728150.
12. Zimmermann P. Exploring the microbial landscape of the nasopharynx in children: a systematic review of studies using next generation sequencing. *Front Microbiomes*. 2023;2:1231271. [10.3389/frmbi.2023.1231271](http://dx.doi.org/10.3389/frmbi.2023.1231271). . PMID: 41853378; PMCID: PMC12993585.
13. Kaur R, Pichichero M, Colonization. Density, and Antibiotic Resistance of *Streptococcus pneumoniae*, *Haemophilus Influenzae*, and *Moraxella catarrhalis* among PCV13-Vaccinated Infants

- in the First Six Months of Life in Rochester, New York: A Cohort Study. *J Pediatric Infect Dis Soc.* 2023;12(3):135–142. 10.1093/jpids/piad004. PMID: 36645216.
14. Deng WJ, Zhang JF, Li PY, Zhou JL, Yao ZJ, Ye XH. Co-carriage of *Streptococcus pneumoniae* and *Moraxella catarrhalis* among preschool children and its influencing factors. *Zhongguo Dang Dai Er Ke Za Zhi.* 2022;24(8):874–80. 10.7499/j.issn.1008-8830.2204163. English, Chinese.
 15. Trusinska D, Zin ST, Sandoval E, Homaira N, Shi T. Risk Factors for Poor Outcomes in Children Hospitalized With Virus-associated Acute Lower Respiratory Infections: A Systematic Review and Meta-analysis. *Pediatr Infect Dis J.* 2024;43(5):467–76. Epub 2024 Jan 26. PMID: 38285519; PMCID: PMC11003409.
 16. Bähre J, Lange M, Salaschek P, Twardella D, Arens S, Eberhard F, Barten-Neiner G, Panning M, Köster H, Körner-Rettberg C, Wetzke M, Happel C. Distinct seasonality and increased respiratory failure in RSV patients < 2 years of age after emergence of SARS-CoV-2: data from the multicentric, prospective PAPI study. *Eur J Pediatr.* 2025;184(4):246. 10.1007/s00431-025-06057-0. . PMID: 40080186; PMCID: PMC11906531.
 17. Pickles RJ, Chen G, Randell SH. Enhanced susceptibility of pediatric airway epithelium to respiratory syncytial virus infection. *J Clin Invest.* 2024;134(21):e185689. 10.1172/JCI185689. . PMID: 39484717; PMCID: PMC11527439.
 18. Oo SWC, Khoo SK, Cox DW, Chidlow G, Franks K, Prastanti F, Bochkov YA, Borland ML, Zhang G, Gern JE, Smith DW, Bizzintino JA, Laing IA, Le Souëf PN. Defining Age-specific Relationships of Respiratory Syncytial Virus and Rhinovirus Species in Hospitalized Children With Acute Wheeze. *Pediatr Infect Dis J.* 2021;40(10):873–879. 10.1097/INF.0000000000003194. PMID: 34321447.
 19. Leung NHL. Transmissibility and transmission of respiratory viruses. *Nat Rev Microbiol.* 2021;19(8):528–45. 10.1038/s41579-021-00535-6. . Epub 2021 Mar 22. PMID: 33753932; PMCID: PMC7982882.
 20. Radzikowska U, Eljaszewicz A, Tan G, Stocker N, Heider A, Westermann P, Steiner S, Dreher A, Wawrzyniak P, Rückert B, Rodriguez-Coira J, Zhakparov D, Huang M, Jakiela B, Sanak M, Moniuszko M, O'Mahony L, Jutel M, Keadze T, Jackson DJ, Edwards MR, Thiel V, Johnston SL, Akdis CA, Sokolowska M. Rhinovirus-induced epithelial RIG-I inflammasome suppresses antiviral immunity and promotes inflammation in asthma and COVID-19. *Nat Commun.* 2023;14(1):2329. 10.1038/s41467-023-37470-4. Erratum in: *Nat Commun.* 2023;14(1):3493. doi: 10.1038/s41467-023-39275-x. PMID: 37087523; PMCID: PMC10122208.
 21. Liu J, Wang W, Cao K, Ren Z, Fu X, Chen Y, Wang W, Zheng Y, Bao Y, Luo X, Chen J. Epidemiology and clinical characteristics of human rhinovirus in hospitalized children and adolescents with acute respiratory infections: a longitudinal study in Shenzhen, China (2019–2024). *Virol J.* 2025;22(1):280. 10.1186/s12985-025-02901-9. . PMID: 40814124; PMCID: PMC12351938.
 22. Hu L, Yang Y, Lin J, Yan Q, Sun C, Li Z, Sun L, Xu J, Chen J, Bai G. Epidemiological characteristics of respiratory syncytial virus infection in pediatric patients before, during the COVID-19 pandemic and

- after easing of COVID-19 restrictive measures in China. *J Med Virol.* 2024;96(1):e29374. 10.1002/jmv.29374. PMID: 38197487.
23. Morelli T, Freeman A, Staples KJ, Wilkinson TMA. Hidden in plain sight: the impact of human rhinovirus infection in adults. *Respir Res.* 2025;26(1):120. 10.1186/s12931-025-03178-w. . PMID: 40155903; PMCID: PMC11954295.
 24. Staadegaard L, Caini S, Wangchuk S, Thapa B, de Almeida WAF, de Carvalho FC, Fasce RA, Bustos P, Kyncl J, Novakova L, Caicedo AB, de Mora Coloma DJ, Meijer A, Hooiveld M, Huang QS, Wood T, Guiomar R, Rodrigues AP, Lee VJM, Ang LW, Cohen C, Moyes J, Larrauri A, Delgado-Sanz C, Demont C, Bangert M, Dückers M, van Summeren J, Paget J. Defining the seasonality of respiratory syncytial virus around the world: National and subnational surveillance data from 12 countries. *Influenza Other Respir Viruses.* 2021;15(6):732–41. 10.1111/irv.12885. . Epub 2021 Jul 13. PMID: 34255934; PMCID: PMC8542954.
 25. Jiang W, Wu M, Zhou J, Wang Y, Hao C, Ji W, Zhang X, Gu W, Shao X. Etiologic spectrum and occurrence of coinfections in children hospitalized with community-acquired pneumonia. *BMC Infect Dis.* 2017;17(1):787. 10.1186/s12879-017-2891-x. . PMID: 29262797; PMCID: PMC5738861.
 26. Wang J, Wang R, Xie Z. Rethinking Paediatric Respiratory Infections: The Role of Mixed Pathogen Infections. *Rev Med Virol.* 2025;35(2):e70021. 10.1002/rmv.70021. PMID: 40000823.
 27. Li ZJ, Zhang HY, Ren LL, Lu QB, Ren X, Zhang CH, Wang YF, Lin SH, Zhang XA, Li J, Zhao SW, Yi ZG, Chen X, Yang ZS, Meng L, Wang XH, Liu YL, Wang X, Cui AL, Lai SJ, Jiang T, Yuan Y, Shi LS, Liu MY, Zhu YL, Zhang AR, Zhang ZJ, Yang Y, Ward MP, Feng LZ, Jing HQ, Huang LY, Xu WB, Chen Y, Wu JG, Yuan ZH, Li MF, Wang Y, Wang LP, Fang LQ, Liu W, Hay SI, Gao GF, Yang WZ, Chinese Centers for Disease Control and Prevention (CDC). Etiology of Respiratory Infection Surveillance Study Team. Etiological and epidemiological features of acute respiratory infections in China. *Nat Commun.* 2021;12(1):5026. 10.1038/s41467-021-25120-6. . PMID: 34408158; PMCID: PMC8373954.
 28. Xu Q, Pichichero ME. Co-colonization by *Haemophilus influenzae* with *Streptococcus pneumoniae* enhances pneumococcal-specific antibody response in young children. *Vaccine.* 2014;32(6):706–11. 10.1016/j.vaccine.2013.11.096. . Epub 2013 Dec 16. PMID: 24355091; PMCID: PMC4407564.
 29. Yu X, Yang X, Song Y, Yu J, Jiang T, Tang H, Yang X, Zeng X, Bi J, Shen A, Sun L. Lower respiratory tract co-infection of *Streptococcus pneumoniae* and respiratory syncytial virus shapes microbial landscape and clinical outcomes in children. *Front Cell Infect Microbiol.* 2025;15:1593053. 10.3389/fcimb.2025.1593053. . PMID: 40552118; PMCID: PMC12183261.
 30. Lin T, Wang Z, Chen W, Pan J, Zhao Y, Zhang Y, Wu H, Liu X, Sun X, Yang Y. Post-pandemic shifts in pediatric respiratory pathogens and co-infection dynamics in Xi'an and surrounding areas: a large-scale retrospective study from 2021 to 2025. *BMC Infect Dis.* 2025;25(1):1751. 10.1186/s12879-025-12128-2. . PMID: 41462117; PMCID: PMC12752169.

Figures

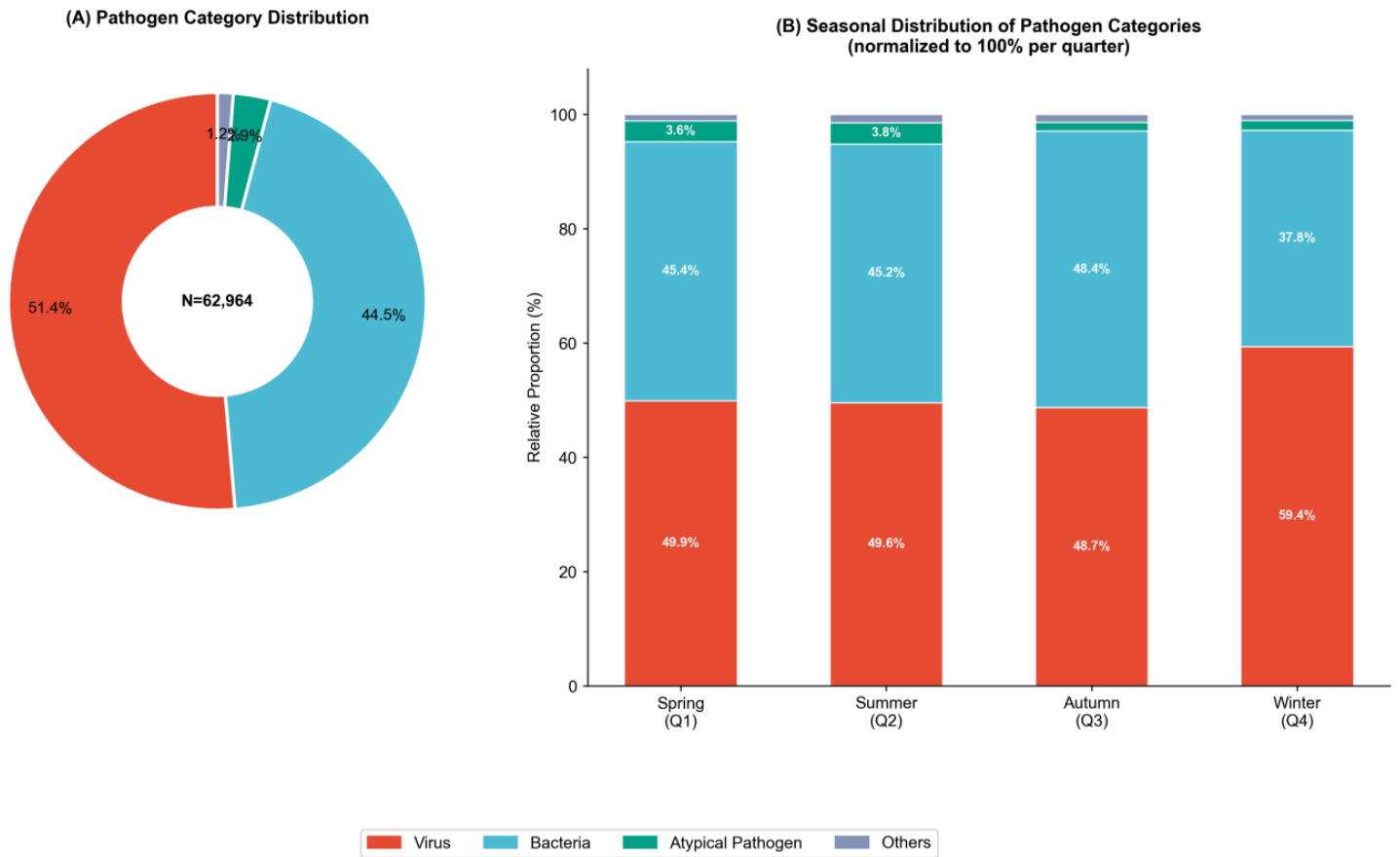


Figure 1

Pathogen Category Composition and Seasonal Distribution.

(A) Donut chart showing the proportion of Bacteria, Virus, Atypical Pathogen, and Others among all positive detections. (B) Seasonally stratified stacked bar chart (normalized to 100% per quarter). N = 28,736 specimens.

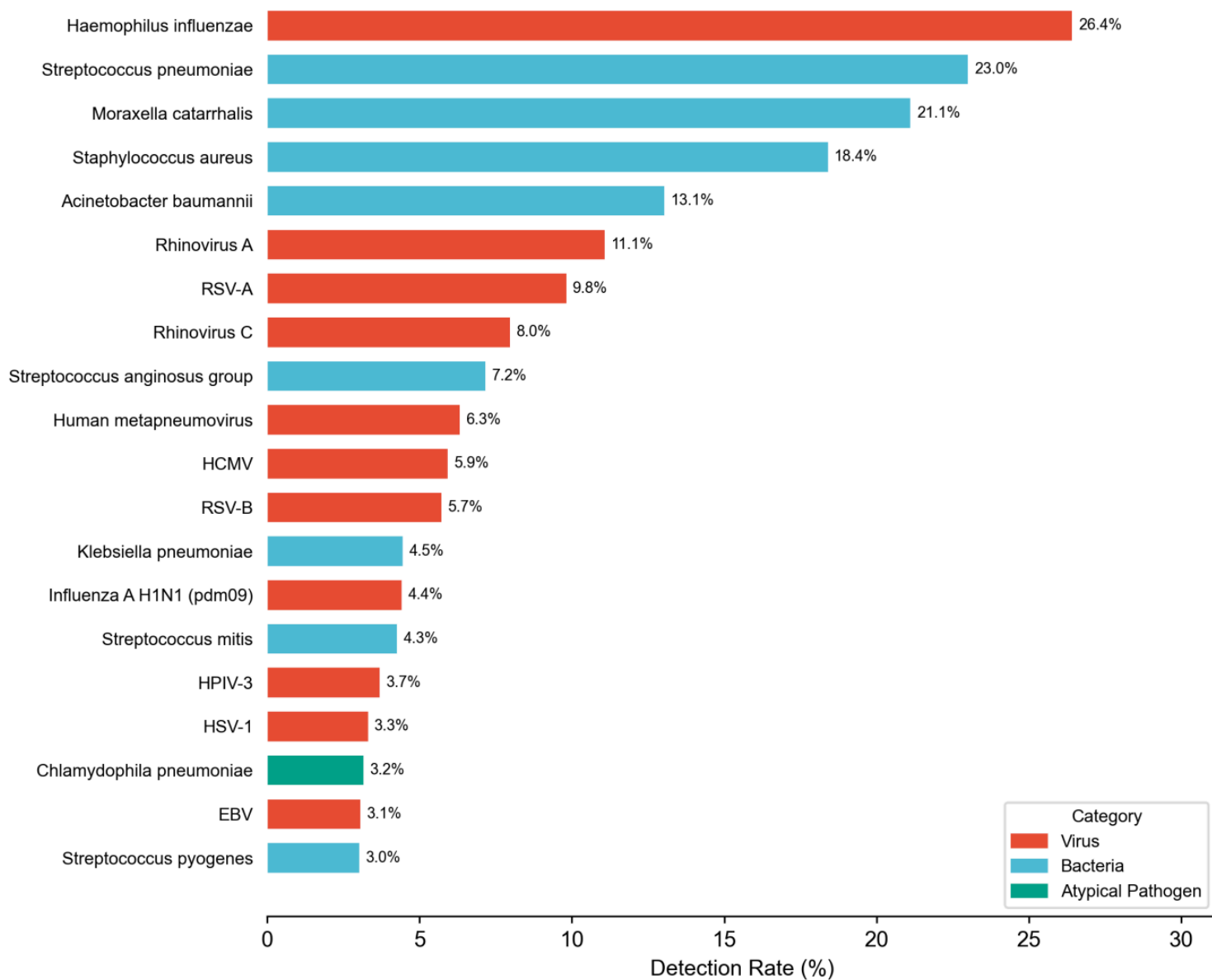


Figure 2

Detection Rates of the Top 20 Respiratory Pathogens.

Horizontal bar chart displaying the top 20 pathogens by detection rate (%). Color coding indicates pathogen category: Bacteria (blue), Virus (orange), Atypical Pathogen (green). N = 28,736.

Figure 3. Age-stratified Detection Rates of Top Pathogens (%)
(Chi-square test, all $p < 0.001^{***}$)

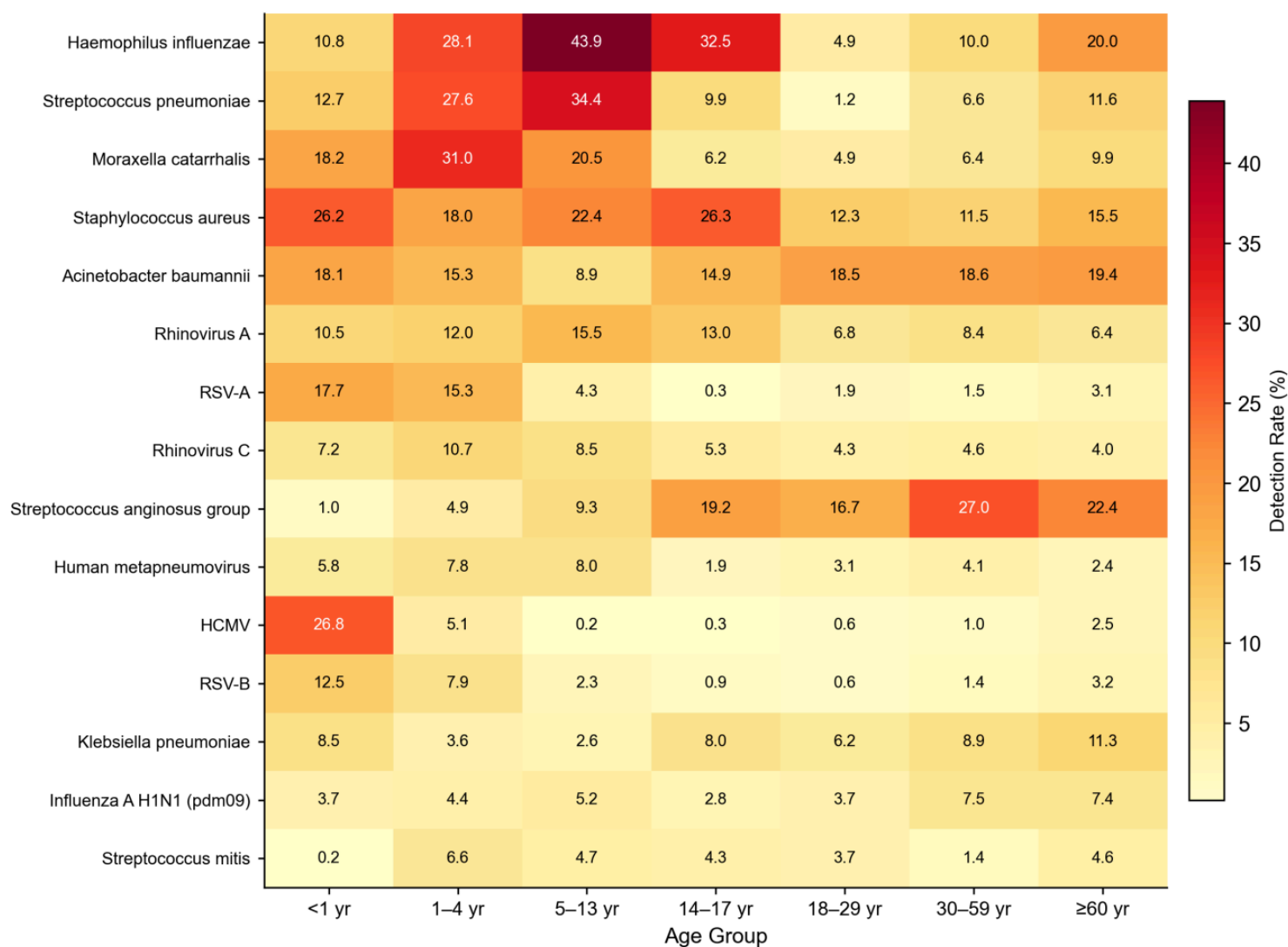


Figure 3

Age-Stratified Detection Rate Heatmap of the Top 15 Pathogens.

Heatmap illustrating the detection rate (%) of the 15 most prevalent pathogens across seven age groups. Color intensity reflects the magnitude of detection rate. N = 28,736 specimens.

Figure 4. Monthly Detection Trends of Top 8 Pathogens

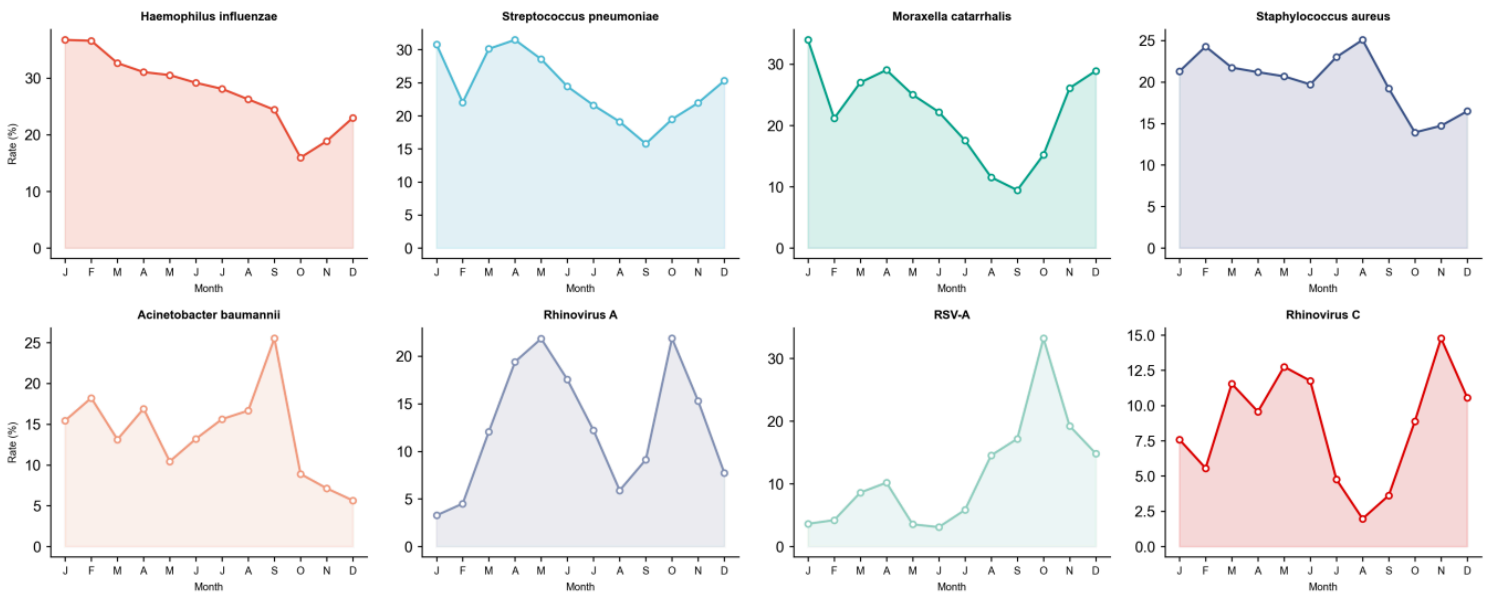


Figure 4

Monthly Trends of the Top 8 Pathogens (January-December 2025).

Line plots showing monthly detection rates (%) of the eight most prevalent pathogens. Rhinovirus A/C peak in Spring and Autumn; RSV-A peaks in Winter (November-December). N = 28,736.

Figure 5. Regional Distribution of Pathogen Categories (normalized to 100% per region)

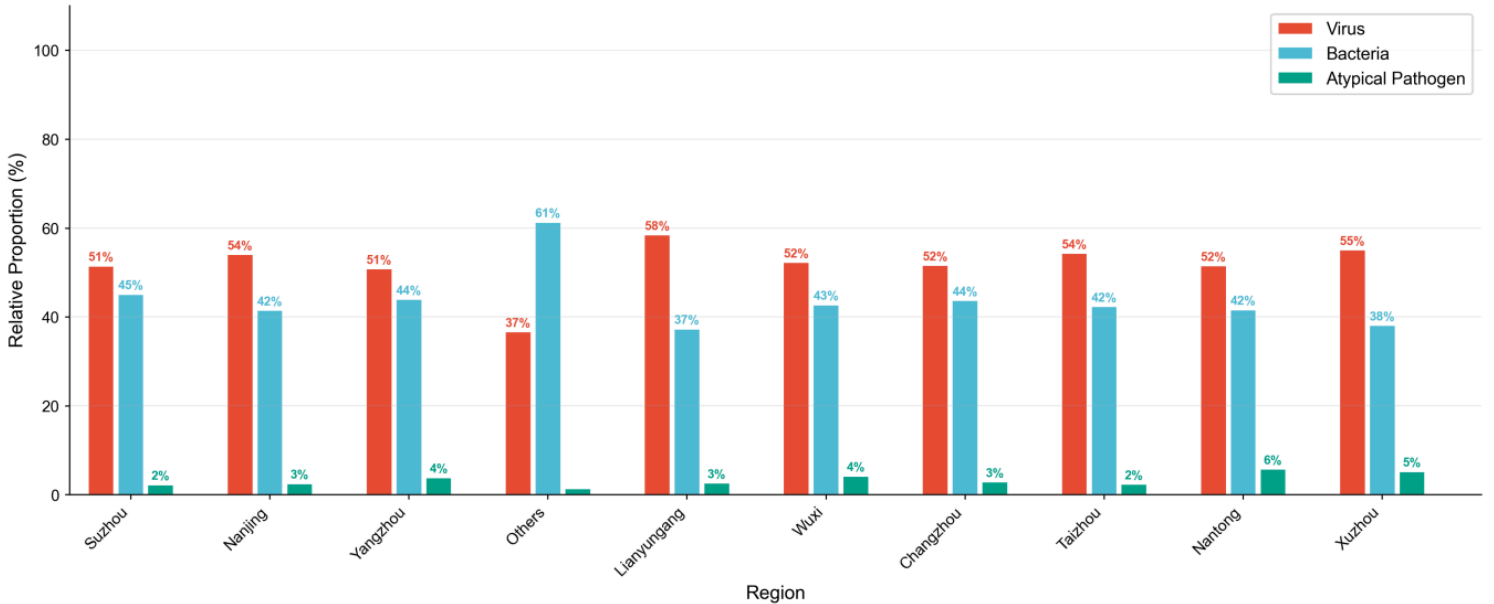


Figure 5

Regional Distribution of Pathogen Categories Across Major Cities.

Grouped bar chart showing the relative proportion (%) of Bacteria, Virus, Atypical Pathogen, and Fungi across the 10 most represented cities. Data are normalized per city. Suzhou: Bacteria 56.4%, Virus 22.7%; Nanjing: Virus 23.2%. N = 28,736.

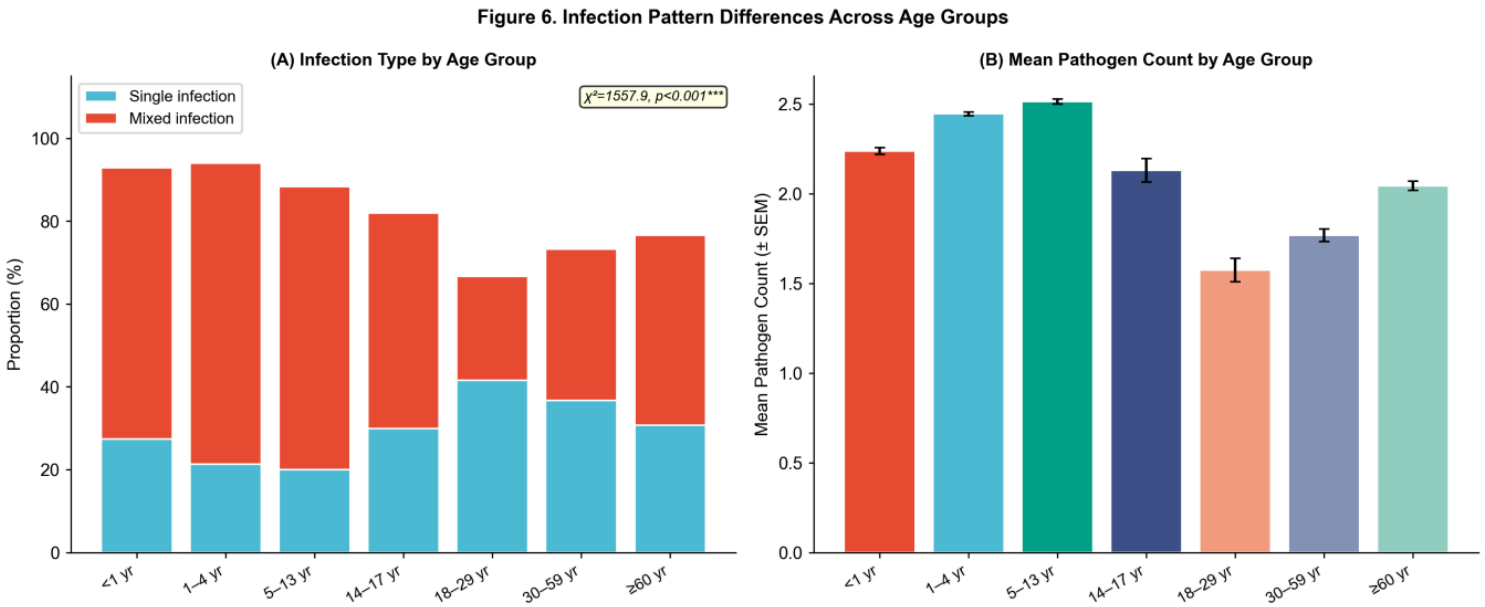


Figure 6

Single vs. Mixed Infection Patterns by Age Group.

(A) Stacked bar chart showing the proportion of single infection (blue) and mixed infection (orange) across age groups. (B) Mean number of pathogen categories detected per specimen by age group. N = 28,736.

Figure 7. Top 20 Pathogen Co-occurrence Pairs in Mixed Infections

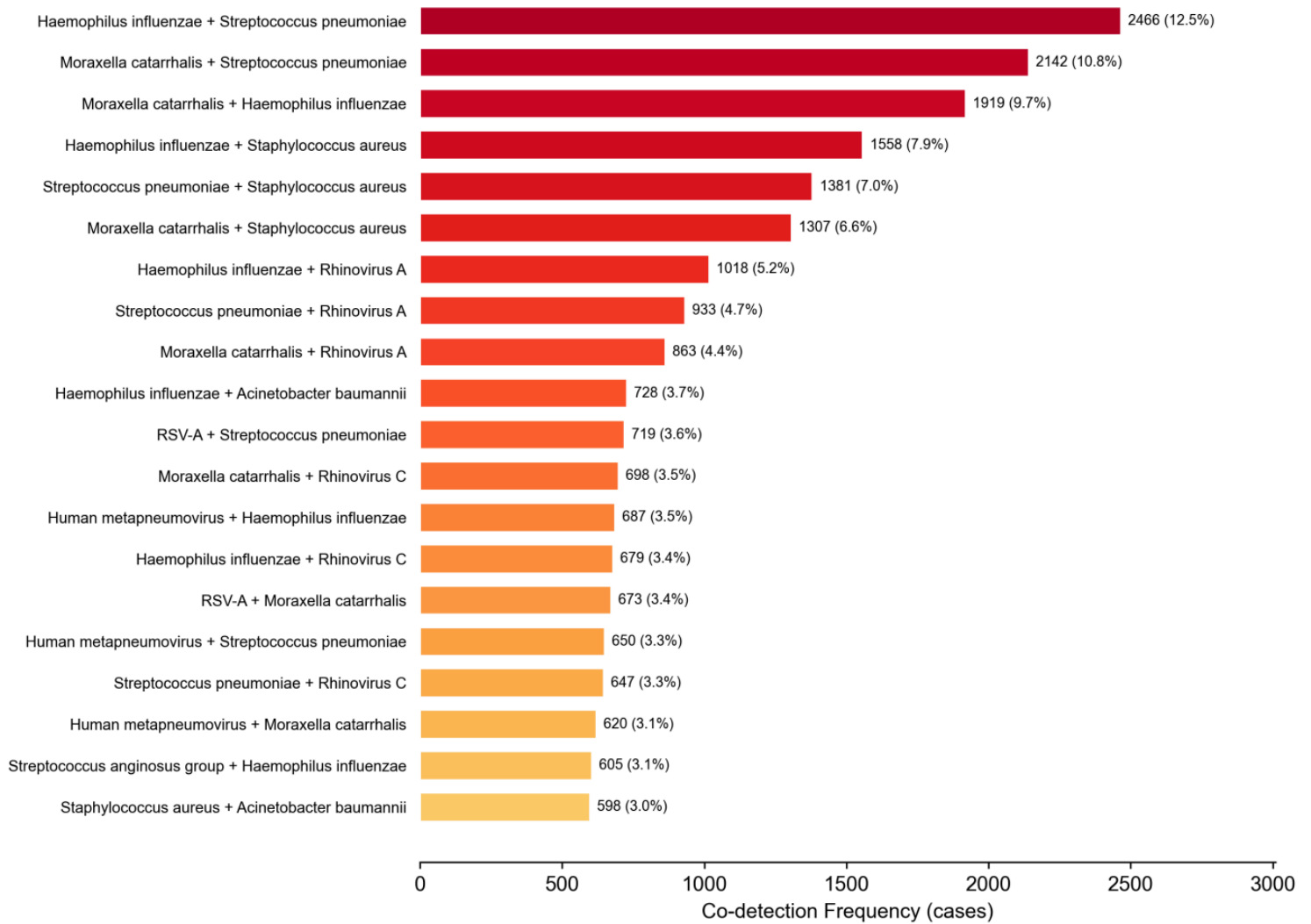


Figure 7

Top 20 Pathogen Co-infection Pairs.

Horizontal bar chart of the 20 most frequently detected co-infection pairs. Color intensity indicates combination type: bacterial-bacterial (dark blue), viral-bacterial (orange), others (gray).

Figure 8. Monthly Detection Rate Heatmap for Top Pathogens (%)

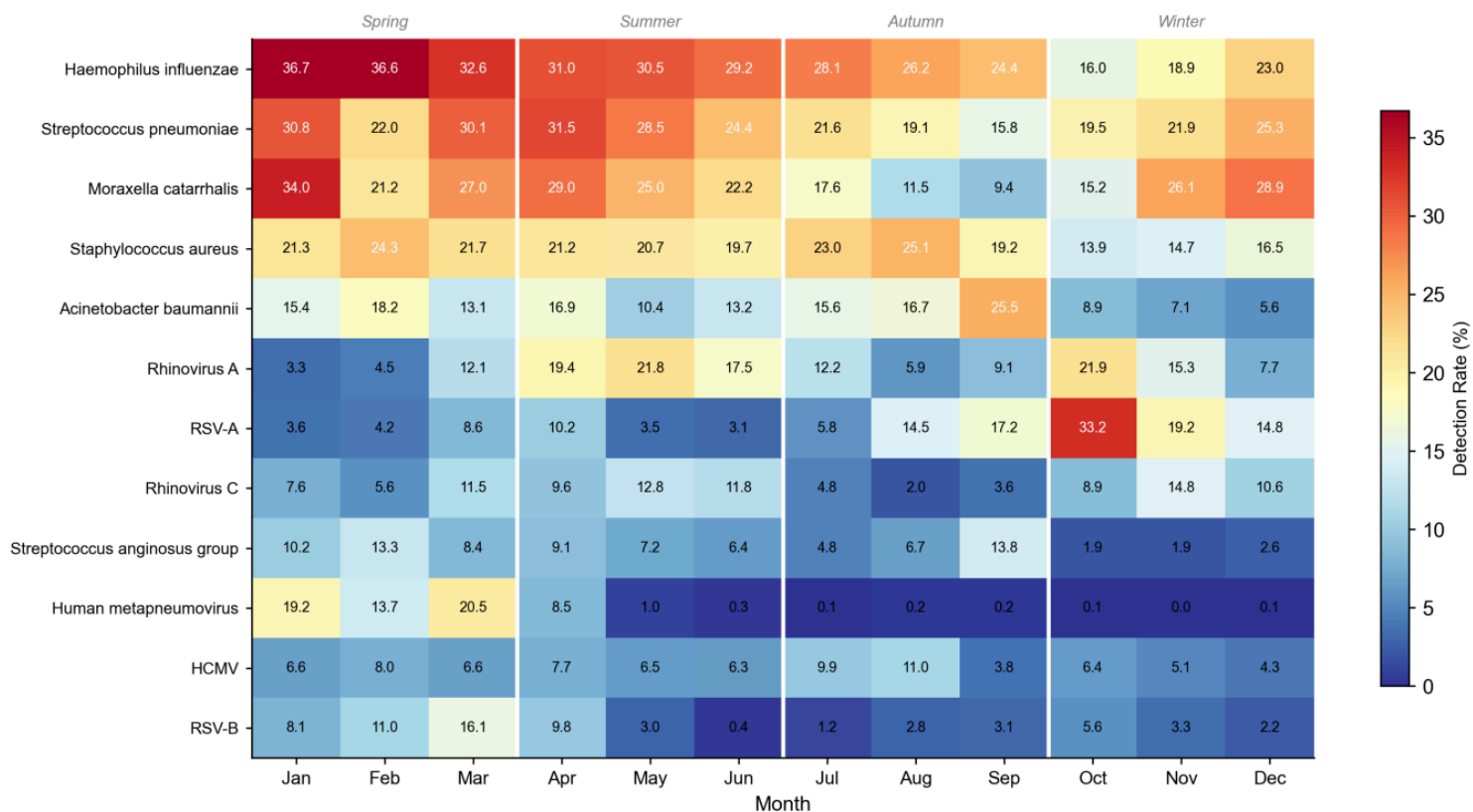


Figure 8

Monthly Detection Rate Heatmap of the Top 12 Pathogens.

Heatmap showing monthly detection rates (%) of the 12 most prevalent pathogens across all 12 months of 2025. Seasonal separators (dashed lines) delineate Spring, Summer, Autumn, and Winter. N = 28,736 specimens.