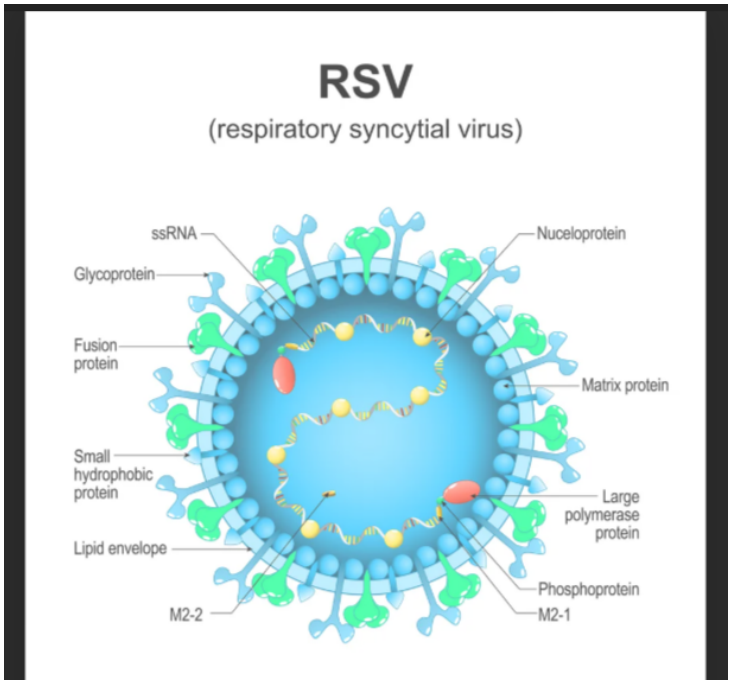
DH 307 - Report

Analysis of RSV Data



Varun Ram Narayanan (22B0347)

1. Introduction

1. **Transmission and Incubation**
   (a) Spread through droplets from the nose and throat of infected individuals.
   (b) Incubation period: 4–6 days.
2. **Symptoms**
   (a) **Mild upper respiratory symptoms:** runny nose, sore throat, headache, fatigue, fever.
   (b) **Severe lower respiratory symptoms:** cough, wheezing, rapid breathing, shortness of breath due to airway obstruction from mucus and dead cells.
   (c) **Extreme cases:** dangerously low oxygen levels, respiratory failure, and death.
3. **Epidemiology**
   (a) Nearly all children contract RSV by age two.
   (b) Severe cases are most common in otherwise healthy infants.
   (c) High-risk groups: premature babies, children with chronic lung or heart disease.
   (d) Early-life RSV infection can lead to long-term consequences such as:
      (i) Recurrent wheeze
      (ii) Hospitalizations for respiratory illness
      (iii) Impaired lung function
   (e) Adults are reinfected throughout life, usually mild in younger adults but potentially severe in the elderly.
4. **RSV Subtypes: A vs B**
   (a) Primary difference: genetic makeup, specifically in the **G protein**, which helps the virus attach to human cells.
   (b) G protein variation:
      (i) RSV A and RSV B G proteins share only ∼53% amino acid sequence.
      (ii) This genetic difference allows classification into RSV A or RSV B.
   (c) **F protein:** much more similar between the two subgroups; main target for vaccines and antibody treatments.

2. Literature Review

2.1 North Indian Pediatric Studies

Agarwal et al. (2023): In a cross-sectional study of children under five with upper respiratory infections in North India, RSV was detected in 23.5% of cases. The majority of these infections were mild and self-limiting, with none requiring hospital admission or leading to severe illness. The study emphasized the predominance of viral causes in URIs, supporting minimal antibiotic use.

> **Key Insight:** RSV infections in India are highly seasonal, primarily affect early childhood, and can range from mild URIs to severe lower respiratory distress requiring ICU care. These studies also indicate widespread but often unnecessary antibiotic use.

Aneja et al. (2024): This study reported RSV in 31% of severe hospitalizations among children, strongly linked to severe acute respiratory infections (SARI). RSV cases peaked during the winter–monsoon season, underscoring the seasonal burden. The findings highlighted RSV as a major contributor to pediatric respiratory morbidity requiring hospitalization.

> **Key Insight:** RSV contributes significantly to hospitalizations in young children and follows strong seasonal trends in North India.

Angurana et al. (2021): Among infants with acute viral bronchiolitis in a tertiary care hospital, RSV was the predominant pathogen accounting for 51% of cases. Notably, one-third of these infants required intensive care admission, and the case fatality rate was 8%, indicating significant morbidity and mortality risk associated with RSV bronchiolitis in infancy.

> **Key Insight:** RSV infections in infancy can range from moderate illness to severe bronchiolitis requiring ICU care, highlighting the clinical significance of early detection and management.

2.2 RSV Trends during COVID-19

Bhardwaj et al. (2022, 2024): During the early COVID-19 pandemic (2020–2021) in Pune, strict public health measures such as masks and social distancing sharply reduced RSV circulation. Once restrictions eased in 2021, RSV cases increased rapidly. The 2024 study showed fluctuating prevalence of RSV-A and RSV-B, persistent dominance in children aged 0–4 years, and no significant gender differences.

Key Insight: COVID-19 measures temporarily suppressed RSV, but the virus resurged as restrictions eased, particularly affecting young children, with fluctuating subtype prevalence.

Kang et al. (2023): RSV positivity in pediatric ALRTIs dropped drastically during the COVID-19 Delta and Omicron waves in northern India, while rhinovirus and bocavirus increased. Pre-pandemic RSV rates were highest, especially in infants, and the pandemic altered typical viral circulation patterns.

Key Insight: COVID-19 disrupted typical RSV seasonality, demonstrating the sensitivity of RSV transmission to public health interventions.

2.3 Eastern Uttar Pradesh

Deval et al. (2024): RSV-B caused 12.65% of severe acute respiratory infections in children under five, mainly during late autumn and winter. Children with RSV-B often presented with fever, cough, breathlessness, and sore throat, with many requiring oxygen supplementation and ICU care. Co-infections occurred in 15% of cases, worsening clinical severity, and mortality among RSV-B patients was notable.

Key Insight: RSV-B infections peak seasonally (October–winter) and can lead to severe clinical outcomes, especially with co-infections.

2.4 Eastern India Outbreak

Ghosh et al. (2022): RSV-A was found in 72.58% of admitted children, mostly infants under one year, with male predominance. Major symptoms included tachypnea, cough, fever, and hypoxia. Approximately 30% required PICU admission, and 21% required mechanical ventilation. Independent predictors for PICU admission were prematurity, age below one year, congenital heart disease, and hypoxia at admission.

Key Insight: RSV-A causes severe illness in infants, emphasizing the importance of early identification, monitoring, and clinical management.

2.5 South India

Kini et al. (2019): RSV in South India peaks after monsoon (August–November), predominantly affects children under 2, and is associated with shorter hospital stays but severe respiratory complications. RSV prevalence is high among infants, with seasonal peaks, and larger outbreaks often follow seasonal patterns. No gender differences were observed.

> **Key Insight:** RSV in South India follows seasonal post-monsoon peaks, mainly affecting infants and young children, with consistent outbreak patterns.

3. RSV Data Dashboard Analysis

3.1 Dataset

The dataset consists of 2169 patient entries with a set of 166 features detailing demographics, geography, symptoms, diseases, attributes of the hospital stay, and various treatments provided for the given patients.

Main observations regarding the dataset:

(a) This data exclusively consists of that of children aged from 0 to 5 years. Almost **96%** of the cases are those of children aged 0 to 1. There is no real correlation with the economic status and the frequency of occurrence of the RSV virus. In fact, the % positive rates were in fact higher for the people above the poverty who presumably have better access to healthcare. However, the number of people tested below the poverty line are significantly higher than those above the poverty line in this dataset.

(b) There are 2 main RSV Variants, RSV A and RSV B while there are also a significant fraction of cases labelled **RSV_A_B** which can be co-infection or unspecified cases. This variant has a strong positive correlation with the RSV B variant (**0.73**), but a weak positive correlation with the RSV A(**0.08**) variant. We have looked at various aspects of the disease to understand what are the main contributing factors or any patterns that seem to be prevalent.

3.2 Analysis

The RSV dashboard provides an interactive interface to analyze Respiratory Syncytial Virus (RSV) data across multiple dimensions. The analyses implemented include the following:

1. **Variant Selection:** Users can select among RSV subtypes (`RSV A`, `RSV B`, or both) or `Deaths` data. Each variant allows exploration of multiple features such as state, region, age, symptoms, seasonality, and vaccine status.

2. **Geographic Analysis:**

   (a) **State-wise Mapping:** State-level RSV positivity rates (% Positive) are visualized on an India map, highlighting the top affected states.

   (b) **Region-wise Analysis:** Pie charts represent the distribution of RSV positivity across regions.

   (c) **Main Observations:**

      (i) The maximum fraction of cases appear in the South and Central India for the RSV A variant, while it is most prevalent in the Northeast, South, and North India for the RSV B variant. The unspecified cases are more prevalent in South and North India.

      (ii) By looking at the data state-wise, we were able to draw additional insights regarding concentration of cases. Firstly, all reported cases from the Northeast came exclusively from the state of Tripura. Cases in the North, West, Central, and East of India are fairly scattered across states. When it comes to the South of India, RSV A cases are fairly scattered while RSV B and unspecified cases are fairly concentrated in the state of Pondicherry.

3. **Symptoms Analysis:**

   (a) **Symptoms vs Region:** Bar charts compare RSV Positive and Negative cases by symptom within each region, highlighting top symptoms.

   (b) **Symptoms vs Variant:** Comparative plots show RSV Positive and Negative prevalence of symptoms across variants, with a separate analysis of the difference in symptom prevalence.

   (c) **Main Observations:**

      (i) There are people coming to hospitals with such a wide variety of symptoms and thus I looked at the symptoms of those who tested positive for RSV against those who did not. For the RSV A variant, the most frequently occurring symptoms in both the populations were Cough, Shortness of Breath, and Respiratory Diseases and there was no strong distinguishing factor.

      (ii) However, there were some symptoms like Conjunctivitis, Ear Pain, Rashes, and Cyanosis which are almost never associated with the virus. Similar patterns were seen across all variants when it comes to these symptoms. Thus, we tried breaking it down and analyzing it on a regional level. However, there was similar uniformity among disease occurrences on a regional level as well.

4. **Seasonal Trends:** RSV cases are analyzed over months and years to visualize seasonal patterns and year-over-year comparison.

(a) **Main Observation:** The main takeaway when we looked at seasonal onset of RSV virus was that a significant fraction of the RSV A variant cases were those detected in October to December 2024. On the other hand, RSV B and the unspecified variant cases were mostly onset in the same timeframe in the year 2023. This is indicative that the virus tends to onset in the winter season.

5. **Vaccine Analysis:** The dashboard compares RSV positivity among individuals based on vaccine status, showing the impact of vaccination on infection rates.

   (a) **Main Observation:** The children who have been given the Child Pentavalent Vaccine (Diphtheria, Pertussis, Tetanus, Hepatitis B, Haemophilus influenzae type b) appear to be less vulnerable to the RSV variant B while there appears to be no significant benefit of taking the vaccination when it comes to the RSV A variant as the % positive rates are similar whether the children have taken the vaccines or not. Similar trends are observed for vaccines against measles and pneumonia.

6. **Mortality Analysis:** For the `Deaths` dataset, bar charts show the count of deaths associated with various symptoms, facilitating identification of high-risk conditions.

   (a) **Main Observation:** There are less than 10 deaths associated with the RSV A and B variants in the given dataset. To be exact, there are 3 deaths associated with RSV A, 5 with RSV B, and 3 with the unspecified variant.

7. **Antibiotics:** While a wide range of antibiotics were used by the people part of this dataset, I have extracted the most frequently used antibiotics among those who have tested positive for the RSV Virus. The most frequently taken antibiotic was **CEFTRIAXONE**.

8. **Data Table:** All analyses are supported by an interactive, filterable data table, allowing users to examine the underlying data in detail.

Overall, the dashboard integrates geographic, demographic, symptom, seasonal, and vaccination data to provide a comprehensive overview of RSV prevalence and impact across India.