# Future Projections of Pneumonia Mortality Across Ages: A Global Perspective

## ABSTRACT

Pneumonia continues to be a major health issue worldwide, especially impacting those at-risk across various age categories. By incorporating comprehensive modeling techniques, this study addresses the pressing need for accurate global pneumonia mortality forecasting across age demographics to inform evidence-based interventions and policymaking. This study employs forecasting models, including Auto-Regressive Integrated Moving Average Models (ARIMA), Generalized Additive Models (GAM), Simple Linear Regression (SLR), Facebook's Prophet Model (Prophet), and the *n*-sub-epidemic framework, to predict pneumonia mortality trends. Data from 1990 through 2019 were used for calibration, with predictive forecasts extending through 2030. The percent change in pneumonia deaths from 2020 to 2030 along with model uncertainty were assessed for all model forecasts. Global trends revealed a consistent rise in pneumonia mortality among individuals aged ≥70 years, a gradual increase in ages 50–69, while those under 5 showed a decline. The *n*-sub-epidemic model predicts a decrease in mortality for ages 15–49 and 5–14, contrasting with statistical models showing increasing trends among ages 15–49. The consistency of forecasting trends across models validates the study's findings and indicates their adaptability to different epidemiological contexts.



**Keywords:** Pneumonia, forecasts, statistical models, *n*-sub-epidemic, global

**Future Projections of Pneumonia Mortality Across Ages:**

**A Global Perspective**

by

Dhanisha Nerurkar

A Thesis Submitted to the Graduate Faculty

of Georgia State University in Partial Fulfillment

of the

Requirements for the Degree

MASTER OF PUBLIC HEALTH

ATLANTA, GEORGIA

30303

**APPROVAL PAGE**

Future Projections of Pneumonia Mortality Across Ages: A Global Perspective

by

Dhanisha Nerurkar

Approved:

Dr. Gerardo Chowell
Committee Chair

Dr. Ruiyan Luo
Committee Member

Amanda Bleichrodt
Committee Member

December 10, 2024
Date

## Author's Statement Page

In presenting this thesis as a partial fulfillment of the requirements for an advanced degree from Georgia State University, I agree that the Library of the University shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to quote from, to copy from, or to publish this thesis may be granted by the author or, in his/her absence, by the professor under whose direction it was written, or in his/her absence, by the Associate Dean, School of Public Health. Such quoting, copying, or publishing must be solely for scholarly purposes and will not involve potential financial gain. It is understood that any copying from or publication of this dissertation which involves potential financial gain will not be allowed without written permission of the author.

Dhanisha Nerurkar

Signature of Author

## TABLE OF CONTENTS

## LIST OF FIGURES

# CHAPTER I – INTRODUCTION

Pneumonia is a respiratory infection characterized by the build-up of fluids and pus in the alveoli. *Streptococcus pneumoniae* is the most common cause of bacteria associated pneumonia [1]. Certain populations are at a higher risk of developing pneumonia due to underlying factors like age, weakened immune systems, and chronic illnesses [2]. Acknowledging these factors and tackling the underlying causes is essential for the effective management and treatment of pneumonia.

Worldwide estimates from the World Health Organization (WHO) show that pneumonia exerts a profound impact as a leading cause of global mortality across the population. Particularly, among children aged 1 to 5 years, it accounted for about 22% of all deaths in 2019 [3]. Additionally in the same year, adults aged 70 years and older experienced the highest rates of pneumonia globally [4]. This highlights the vulnerability of both young children and the elderly to pneumonia, while mortality rates for other age groups have remained relatively stable over the past decades [5]. Research extensively covers pneumonia burden in children under 5, but data on adults and older children is limited [4].

Pneumonia stands as a significant public health concern, causing about 2.5 million deaths across all age groups [5]. The economic burden of pneumonia is substantial, manifested through increased healthcare expenditures, lost productivity, and is a strain on healthcare systems [6-7]. This issue is particularly prominent in middle income countries, which carry a 2-fold greater burden of pneumonia than higher income countries [7]. As a result, individuals in deprived socioeconomic conditions are at a higher risk of pneumonia [8-12]. The various determinants of pneumonia include education, financial status, age, sex, ethnicity, smoking, and environmental pollutants [13-15]. Thus, it vital to address pneumonia not just at the clinical front but also at various socioeconomic levels.

International health efforts, including the Global Action Plan for the Prevention and Control of Pneumonia (GAPP) launched by WHO and UNICEF, aim to decrease pneumonia-related deaths among children [16]. Additionally, initiatives such as Every Breath Counts and the Stop Pneumonia Initiative aim to implement impactful strategies for combatting pneumonia among vulnerable populations [17]. These global health initiatives and the changing epidemiology of lower respiratory infection and pneumonia related mortality [18] emphasizes the importance of accurate mortality forecasting for shaping effective public health strategies, healthcare resource allocation, and policy-making in pneumonia control.

Forecasting pneumonia to predict future morbidity and mortality of the disease is crucial as pneumonia still stands as the leading cause of death and disability globally [19,7]. Despite efforts to mitigate pneumonia at both local and global levels, there still exists a significant burden of disease in developing countries [20,7]. This necessitates the need for large-scale, diverse, and accurate forecasting of pneumonia mortality to inform effective public health strategies, improve healthcare resource allocation, and enhance policies for pneumonia control. Numerous predictions models have been used in studies to assess their use in clinical settings to mitigate or support in early diagnosis of pneumonia [21-24]. Many different forecasting methodologies, including deep machine learning frameworks [25-26], time-series [27], and mathematical models [28], have been used in pneumonia modeling. Infectious disease modeling has expanded in varied disease areas such as COVID-19 [29], Zika [30], mpox [31] and influenza [32]. Regardless of the availability of this multi-faceted forecasting data, limited research addresses predictive analysis of pneumonia mortality across age groups, especially at a global scale, urging further study.

Our research aims to advance pneumonia mortality forecasting methods by leveraging on comprehensive modeling approach. Specifically, we seek to utilize models such as Auto-Regressive Integrated Moving Average Models (ARIMA), Generalized Additive Models (GAM), Simple Linear Regression (SLR), Facebook's Prophet Model (Prophet), and the *n*-sub-epidemic modeling framework to forecast global trends in pneumonia and lower respiratory diseases across various age groups. These forecasts can empower policymakers to implement evidence-based interventions and reduce the burden of lower respiratory infections and preventable pneumonia-related deaths. Through this approach, we aim to contribute to the field of global pneumonia control by providing diverse and nuanced forecasts that inform targeted interventions and public health strategies. It is imperative to acknowledge that addressing pneumonia deaths necessitates tailored approaches specific to different age groups, and our study is poised to make a substantial contribution in this regard.

## CHAPTER II – LITERATURE REVIEW

Pneumonia is a respiratory infection marked by the accumulation of fluid and pus in the alveoli. This inflammatory condition can result from various infectious agents, including bacteria, viruses, or fungi. Among bacterial pathogens, *Streptococcus pneumoniae* is the most prevalent, especially in children [1,3]. Following closely is *Haemophilus influenzae* type b (Hib), which is the second most common bacterium responsible for pneumonia. The etiology of pneumonia varies, and its manifestations can vary from mild to severe across different age groups. Typical signs include cough, difficulty breathing, chest pain, and fever, along with the possibility of life-threatening complications [2,1].

Certain populations face a higher risk due to various underlying factors. For instance, age, compromised immune systems, chronic diseases, and exposure to environmental factors significantly increase vulnerability of certain population [2,5,1]. Malnutrition is found to significantly elevate pneumonia mortality rates among children, while air pollution exacerbates the condition in the elderly [5,1]. Recognizing these risk factors and addressing their underlying causes are essential for effective pneumonia management and treatment.

Pneumonia remains a major global health challenge, with the WHO estimating it to be a leading cause of death, especially among children aged 1 to 5 years, where it contributes to approximately 22% of all fatalities [3]. Additionally, data from 2019 revealed that pneumonia was most prevalent among adults aged 70 and older [4], highlighting increased susceptibility in both the elderly and young children. In contrast, mortality rates for those between 15 and 69 years and individuals over 5 years, have remained fairly consistent over the years [5].

In 2021, pneumonia caused around half a million deaths among children; this was a notable reduction of about a 28% drop in childhood pneumonia deaths from 2019. The observed decrease in the number of pneumonia deaths among children was considered as the largest reduction in three years by the Global Burden of Disease (GBD) since 1990. The decrease is largely attributed to the impact of COVID-19 restrictions, highlighting the success of preventive strategies in reducing the transmission of infections [20,33]. Additionally, lower respiratory infections ranked as the 5th leading cause of disease burden globally in 2021, with a projection to move to the 9th position by 2050 due to expected improvements in disease management [19]. Despite these improvements in overall disease burden, lower respiratory tract infections continue to pose a significant challenge in low-income and middle-income countries [20,7,34]. Such conditions disproportionately affect individuals from disadvantaged socioeconomic backgrounds, increasing their risk of contracting pneumonia. [8-12]. Research indicates that older adults aged 65 and over experience higher rates of mortality and hospitalization due to pneumonia, with 3.1 million hospitalizations recorded in industrialized countries and 3.7 million in developing nations [35-36], resulting in considerable economic burden associated with pneumonia.

This forthcoming shift in the epidemiology of pneumonia and other lower respiratory infections necessitates large-scale, diverse, and accurate forecasting of pneumonia mortality to inform effective public health strategies, improve healthcare resource allocation, and enhance policies for pneumonia control. Pneumonia mortality forecasting is essential as it enables researchers to predict the spread of outbreaks, evaluate potential interventions, and guide public health policies at a population or community level. Over the years, a variety of forecasting models and frameworks have been utilized for accurate pneumonia diagnosis and treatment. Recent research has concentrated on identifying reliable predictive signs, symptoms, and biomarkers to differentiate pneumonia from other conditions. For example, a prospective observational cohort study by Groeneveld G.H. et al. (2019), conducted from March 2012 to May 2016, evaluated the predictive value of biomarkers for pneumonia diagnosis among patients. They utilized a predictive clinical model which incorporated these predictive factors to diagnose pneumonia. The clinical model found that while CRP emerged as the most effective biomarker, its addition to a clinical model did not significantly enhance prediction accuracy but was useful in guiding antibiotic treatment decisions [22].

Deep learning algorithms and artificial intelligence (AI) have also shown potential in diagnosing pneumonia. Ferreira et al. (2020), Helena Liz et al. (2021), and Ayan E et al. (2022) addressed critical challenges in pneumonia diagnosis using Convolutional Neural Network models and by applying advanced machine learning techniques [25,37,23]. Similarly, Zang Q et al. (2022), proposed a combined GSHHO-FKNN model, which effectively discriminated influenza pneumonia from community-acquired pneumonia, leading to accurate bacterial identification and more effective antibiotic use [26].

In the efforts to assess the efficacy of various existing models in predicting pneumonia, Effah et al. (2022), evaluated the performance of eight machine learning models. Their study found no adequate discriminatory power between individual clinical history, laboratory indicators, and symptoms in diagnosing pneumonia. However, they concluded that ensemble machine learning models, such as Random Forest (RF) and Extreme Gradient Boosting (XGBoost), showed promise in improving prediction accuracy [21].

Another study conducted by Ni Kadek Ary Indah et al. (2022), forecasted the number of under-five pneumonia deaths while comparing the efficacy of ARIMA and Multi-Layer Perceptron (MLP) models. This study employed data spanning from January 2014 to December 2021 to forecast pneumonia mortality trends from January 2022 to December 2022. The study found that the MLP model outperformed ARIMA, achieving a low error percentage with prediction accuracy of 1.27%. This long-term forecasting study for 2022 indicated a peak in deaths during October and an overall increase in mortality compared to the previous year. The model’s superior performance in predicting under-five pneumonia mortality rates underscores its potential utility in shaping public health policies and improving mortality prevention strategies [28].

Other advancements in forecasting methodologies have significantly enhanced our understanding of pneumonia mortality trends and projections to some extent. For instance, Yun JW & Son M (2016) employed a modified age-period-cohort model, using the Nordpred software in R, to forecast cause-specific mortality rates in Korea from 2013 to 2032 based on historical data from 1983 to 2012. The model combined age, period, and cohort effects to enhance the accuracy of mortality forecasts and effectively managed both population dynamics and disease mortality trends. The projections indicated a significant increase in pneumonia-related deaths, forecasting an increase of 319.9% for males and 274.1% for females from 2008-2012 to 2028-2032, largely attributable to the aging population in Korea [38]. An alternative study with similar focus used the ARIMA model to analyze health data related to the leading causes of death, morbidity, and infant mortality in the Philippines [39]. The study, extending projections to 2030, found a general decrease in mortality rates for pneumonia, diabetes mellitus, nephritis, nephrotic syndrome, and some perinatal conditions. Specifically, pneumonia-related mortality among infants was projected to show a slight decreasing trend. Conversely, conditions like bacterial sepsis and respiratory distress were expected to increase. These findings underscore the value of forecasting models in forecasting health trends and the importance of incorporating demographic shifts into public health strategies to address the growing challenge of pneumonia and other age-related diseases.

Despite advancements in forecasting and modeling for diagnosing, early detection, and predicting pneumonia outcomes, there remain significant gaps in delivering comprehensive, age-specific pneumonia mortality forecasts globally. Traditionally, the ARIMA model has been employed to forecast financial trends [40] and weather patterns [41-42], and it has established itself as a benchmark in disease forecasting [43,29,27,44-46]. Generalized Additive Models (GAM), on the other hand, incorporate a set of smooth functions to more effectively capture the influences of various variables, allowing them to identify non-linear trends and thus enhance the performance of conventional generalized linear models [47]. This feature enables them to effectively capture non-linear trends while retaining a level of simplicity and explainability comparable to generalized linear models [48]. Meanwhile Facebook's Prophet model [49], which was initially created for business forecasting, has found increasing use across diverse fields, particularly in infectious disease modeling. It has generated forecasts for conditions such as COVID-19 [50,45,51], mpox [31] and pneumonia forecasts [27]. The *n*-sub-epidemic framework and related ensemble model have also been applied to forecast various infectious diseases and have shown strong performance. [31,29,52,30,32]. During the COVID-19 pandemic Chowell G et al., 2022 [29] employed the *n*-sub-epidemic modelling framework to effectively forecast the COVID-19 trends in the USA. The best performing sub-epidemic and ensemble models demonstrated superior performance in short-term forecasts compared to traditional ARIMA models. Additionally, Cramer EY et al., 2022 [52] also highlighted the importance of multi-model ensemble forecasting in short-term probabilistic forecasting during the COVID-19 pandemic. Their study provided the most accurate forecasting related to COVID-19 deaths at both state and national level. Both these studies accentuate the value of advanced ensemble modeling techniques in improving the accuracy of disease forecasting.

Therefore, this study aims to add to the current body of literature on pneumonia mortality by advancing forecasting methods through the application of comprehensive modeling approach. This study utilizes commonly used statistical models such as ARIMA, GAM, SLR and Facebook's Prophet Model, along with the mathematical framework associated to the *n*-sub-epidemic model, to predict global pneumonia trends across various age groups.

## REFERENCES


1. **National Institute of Health** (2022) What Is Pneumonia? Available at https://www.nhlbi.nih.gov/health/pneumonia (accessed January, 14 2024).
2. **Centers for Disease Control and Prevention** (2022) Pneumonia. Available at https://www.cdc.gov/pneumonia/index.html (accessed January 14 2024).
3. **World Health Organization** (2022) Pneumonia in children. Available at https://www.who.int/news-room/fact-sheets/detail/pneumonia (accessed January 15 2024).
4. **GBD Lower Respiratory Infections Collaborators** (2022) Age-sex differences in the global burden of lower respiratory infections and risk factors, 1990-2019: results from the Global Burden of Disease Study 2019. *Lancet Infect Dis*; **22**: 1626-1647. https://doi.org/10.1016/S1473-3099(22)00510-2.
5. **Dadonaite B. and Roser M.** (2019) Pneumonia. Available at https://ourworldindata.org/pneumonia (accessed January 15 2024).
6. **Broulette J, et al.** (2013) The incidence rate and economic burden of community-acquired pneumonia in a working-age population. *Am Health Drug Benefits*; **6**: 494-503.
7. **JustActions** (2021) The Missing Piece Why the Global Pandemic is an Infection Point for Pnuemonia Control. Revised and updated for COVID-19 and the Global Burden of Disease 2019. Available at https://stoppneumonia.org/wp-content/uploads/2021/10/JUST-ACTIONS_web-spread.pdf (accessed January 15 2024).
8. **Azab SF, et al.** (2014) Impact of the socioeconomic status on the severity and outcome of community-acquired pneumonia among Egyptian children: a cohort study. *Infect Dis Poverty*; **3**: 14. https://doi.org/10.1186/2049-9957-3-14.
9. **Bazie GW, Seid N, Admassu B** (2020) Determinants of community acquired pneumonia among 2 to 59 months of age children in Northeast Ethiopia: a case-control study. *Pneumonia (Nathan)*; **12**: 14. https://doi.org/10.1186/s41479-020-00077-0.
10. **Calvillo-King L, et al.** (2013) Impact of social factors on risk of readmission or mortality in pneumonia and heart failure: systematic review. *J Gen Intern Med*; **28**: 269-282. https://doi.org/10.1007/s11606-012-2235-x.
11. **Cashat-Cruz M, Morales-Aguirre JJ, Mendoza-Azpiri M** (2005) Respiratory tract infections in children in developing countries. *Semin Pediatr Infect Dis*; **16**: 84-92. https://doi.org/10.1053/j.spid.2005.12.005.
12. **Thorn LK, et al.** (2011) Pneumonia and poverty: a prospective population-based study among children in Brazil. *BMC Infect Dis*; **11**: 180. https://doi.org/10.1186/1471-2334-11-180.
13. **Loeb MB** (2004) Use of a broader determinants of health model for community-acquired pneumonia in seniors. *Clin Infect Dis*; **38**: 1293-1297. https://doi.org/10.1086/383469.
14. **Saunders M, Nellums L** (2022) The indirect effects of COVID-19 upon global childhood pneumonia. *Public Health Pract (Oxf)*; **3**: 100228. https://doi.org/10.1016/j.puhip.2022.100228.
15. **Seramo RK, et al.** (2022) Determinants of pneumonia among children attending public health facilities in Worabe town. *Sci Rep*; **12**: 6175. https://doi.org/10.1038/s41598-022-10194-z.
16. **World Health Organization** (2009, June 17) Global Action Plan for Prevention and Control of Pneumonia (GAPP). Available at https://www.who.int/publications/i/item/WHO-FCH-CAH-NCH-09.04 (accessed January 14 2024).
17. **Every Breath Counts** Every Breath Counts Coalition. Available at https://stoppneumonia.org/ (accessed January 14 2024).
18. **GBD Lower Respiratory Infections Collaborators** (2018) Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory infections

in 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Infect Dis*; **18**: 1191-1210. https://doi.org/10.1016/S1473-3099(18)30310-4.
19. **Institute for Health Metrics and Evaluation (IHME)** (2021) Global Burden of Disease 2021: Findings from the GBD 2021 Study., 2024. Seattle, WA. Available at https://www.healthdata.org/research-analysis/library/global-burden-disease-2021-findings-gbd-2021-study (accessed.
20. **GBD LRIs & Antimicrobial Resistance Collaborators** (2024) Global, regional, and national incidence and mortality burden of non-COVID-19 lower respiratory infections and aetiologies, 1990-2021: a systematic analysis from the Global Burden of Disease Study 2021. *Lancet Infect Dis*; **24**: 974-1002. https://doi.org/10.1016/S1473-3099(24)00176-2.
21. **Effah CY, et al.** (2022) Machine learning-assisted prediction of pneumonia based on non-invasive measures. *Front Public Health*; **10**: 938801. https://doi.org/10.3389/fpubh.2022.938801.
22. **Groeneveld GH, et al.** (2019) Prediction model for pneumonia in primary care patients with an acute respiratory tract infection: role of symptoms, signs, and biomarkers. *BMC Infect Dis*; **19**: 976. https://doi.org/10.1186/s12879-019-4611-1.
23. **Liz H, et al.** (2021) Ensembles of Convolutional Neural Network models for pediatric pneumonia diagnosis. *Future Generation Computer Systems*; **122**: 220-233. https://doi.org/10.1016/j.future.2021.04.007.
24. **Schierenberg A, et al.** (2016) External Validation of Prediction Models for Pneumonia in Primary Care Patients with Lower Respiratory Tract Infection: An Individual Patient Data Meta-Analysis. *PLoS One*; **11**: e0149895. https://doi.org/10.1371/journal.pone.0149895.
25. **Ayan E, Karabulut B, Unver HM** (2022) Diagnosis of Pediatric Pneumonia with Ensemble of Deep Convolutional Neural Networks in Chest X-Ray Images. *Arab J Sci Eng*; **47**: 2123-2139. https://doi.org/10.1007/s13369-021-06127-z.
26. **Zhang Q, et al.** (2022) A machine learning framework for identifying influenza pneumonia from bacterial pneumonia for medical decision making. *Journal of Computational Science*; **65**: 101871. https://doi.org/https://doi.org/10.1016/j.jocs.2022.101871.
27. **Hussein HA, Mahmoud MM, Eisa HA** (2023) Performance Evaluation of ARIMA and FB-Prophet Forecasting Methods in the Context of Endemic Diseases: A Case Study of Gedaref State in Sudan. *EAI Endorsed Transactions on Smart Cities*; **7**: e1-e1.
28. **Ni Kadek Ary Indah S, Oka S, Ayu W** (2022) Forecasting Pneumonia Toddler Mortality Using Comparative Model ARIMA and Multilayer Perceptron. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*; **6**: 528-537. https://doi.org/10.29207/resti.v6i4.4106.
29. **Chowell G, et al.** (2022) An ensemble n-sub-epidemic modeling framework for short-term forecasting epidemic trajectories: Application to the COVID-19 pandemic in the USA. *PLoS Comput Biol*; **18**: e1010602. https://doi.org/10.1371/journal.pcbi.1010602.
30. **Oidtman RJ, et al.** (2021) Trade-offs between individual and ensemble forecasts of an emerging infectious disease. *Nature communications*; **12**: 5379.
31. **Bleichrodt A, et al.** (2024) Evaluating the forecasting performance of ensemble sub-epidemic frameworks and other time series models for the 2022–2023 mpox epidemic. *Royal Society Open Science*; **11**. https://doi.org/10.1098/rsos.240248.
32. **Reich NG, et al.** (2019) Accuracy of real-time multi-model ensemble forecasts for seasonal influenza in the US. *PLoS computational biology*; **15**: e1007486.
33. **McAllister DA, et al.** (2019) Global, regional, and national estimates of pneumonia morbidity and mortality in children younger than 5 years between 2000 and 2015: a systematic analysis. *The Lancet Global Health*; **7**: e47-e57. https://doi.org/10.1016/s2214-109x(18)30408-x.

34. **Marangu D, Zar HJ** (2019) Childhood pneumonia in low-and-middle-income countries: An update. *Paediatr Respir Rev*; **32**: 3-9. https://doi.org/10.1016/j.prrv.2019.06.001.
35. **Hespanhol V, Barbara C** (2020) Pneumonia mortality, comorbidities matter? *Pulmonology*; **26**: 123-129. https://doi.org/10.1016/j.pulmoe.2019.10.003.
36. **Shi T, et al.** (2020) Global and Regional Burden of Hospital Admissions for Pneumonia in Older Adults: A Systematic Review and Meta-Analysis. *J Infect Dis*; **222**: S570-S576. https://doi.org/10.1093/infdis/jiz053.
37. **Ferreira JR, et al.** (2020) Multi-View Ensemble Convolutional Neural Network to Improve Classification of Pneumonia in Low Contrast Chest X-Ray Images. *Annu Int Conf IEEE Eng Med Biol Soc*; **2020**: 1238-1241. https://doi.org/10.1109/EMBC44109.2020.9176517.
38. **Yun J-W, Son M** (2016) Forecasting Cause-Specific Mortality in Korea up to Year 2032. *Journal of Korean Medical Science*; **31**: 1181. https://doi.org/10.3346/jkms.2016.31.8.1181.
39. **Talirongan F, Talirongan H, Orong MY.** (2021) Modeling National Trends on Health in the Philippines Using ARIMA. *Journal of Health & Medical Informatics*; **11**: arXiv:2101.01392. https://doi.org/https://doi.org/10.48550/arXiv.2101.01392.
40. **Mondal P, Shit L, Goswami S.** (2014) Study of effectiveness of time series modeling (ARIMA) in forecasting stock prices. *International Journal of Computer Science, Engineering and Applications*; **4**: 13.
41. **Dimri T, Ahmad S, Sharif M.** (2020) Time series analysis of climate variables using seasonal ARIMA approach. *Journal of Earth System Science*; **129**: 1-16.
42. **Tektaş M** (2010) Weather forecasting using ANFIS and ARIMA models. *Environmental Research, Engineering and Management*; **51**: 5-10.
43. **Bleichrodt A, et al.** (2023) Retrospective evaluation of short-term forecast performance of ensemble sub-epidemic frameworks and other time-series models: The 2022-2023 mpox outbreak across multiple geographical scales, July 14(th), 2022, through February 26th, 2023. *medRxiv*. https://doi.org/10.1101/2023.05.15.23289989.
44. **Hyndman** (n.d.) Auto.arima: Fit best ARIMA model to univariate time series. Available at https://www.rdocumentation.org/packages/forecast/versions/8.21.1/topics/auto.arima (accessed).
45. **Long B, Tan, F, Newman M.** (2023) Forecasting the monkeypox outbreak using ARIMA, prophet, NeuralProphet, and LSTM models in the United States. *Forecasting*; **5**: 127-137.
46. **Ruchiraset A, Tantrakarnapa K** (2018) Time series modeling of pneumonia admissions and its association with air pollution and climate variables in Chiang Mai Province, Thailand. *Environ Sci Pollut Res Int*; **25**: 33277-33285. https://doi.org/10.1007/s11356-018-3284-4.
47. **Shafi A** (2021) What are Generalised Additive Models? Available at https://towardsdatascience.com/generalised-additive-models-6dfbedf1350a (accessed 14 April 2024).
48. **Wood SN.** (n.d.) GAM: Generalized additive models with integrated smoothness estimation. Available at https://www.rdocumentation.org/packages/mgcv/versions/1.9-1/topics/gam. (accessed).
49. **Battineni G, Chintalapudi N, Amenta F** (2020) Forecasting of COVID-19 epidemic size in four high hitting nations (USA, Brazil, India and Russia) by Fb-Prophet machine learning model. *Applied Computing and Informatics*.
50. **Kirpich A, et al.** (2022) Excess mortality in Belarus during the COVID-19 pandemic as the case study of a country with limited non-pharmaceutical interventions and limited reporting. *Scientific reports*; **12**: 5475.

51. **Taylor S** PROPHET: Prophet forecaster. Available at https://www.rdocumentation.org/packages/prophet/versions/1.0/topics/prophet (accessed).
52. **Cramer EY, et al.** (2022) Evaluation of individual and ensemble probabilistic forecasts of COVID-19 mortality in the United States. *Proceedings of the National Academy of Sciences*; **119**: e2113561119.

## CHAPTER III – MANUSCRIPT

**Future Projections of Pneumonia Mortality Across Ages: A Global Perspective**

**Dhanisha Nerurkar[1]*, Amanda M Bleichrodt[1], Ruiyan Luo[1], Gerardo Chowell[1,2]***

[1]Department of Population Health Sciences, School of Public Health, Georgia State University, Atlanta, GA, USA.

[2] Department of Applied Mathematics, Kyung Hee University, Yongin 17104, Korea

***Corresponding authors:**

**Name:** Dhanisha Nerurkar, **Email:** dhani.1406@gmail.com

**Name:** Gerardo Chowell, **Email:** gchowell@gsu.edu

**Word Count:** 3963

**Keywords:** Pneumonia, forecasts, statistical models, *n*-sub-epidemic, global

**Abstract**

Pneumonia continues to be a major health issue worldwide, especially impacting those at-risk across various age categories. By incorporating comprehensive modeling techniques, this study addresses the pressing need for accurate global pneumonia mortality forecasting across age demographics to inform evidence-based interventions and policymaking. This study employs forecasting models, including Auto-Regressive Integrated Moving Average Models (ARIMA), Generalized Additive Models (GAM), Simple Linear Regression (SLR), Facebook's Prophet Model (Prophet), and the *n*-sub-epidemic framework, to predict pneumonia mortality trends. Data from 1990 through 2019 were used for calibration, with predictive forecasts extending through 2030. The percent change in pneumonia deaths from 2020 to 2030 along with model uncertainty were assessed for all model forecasts. Global trends revealed a consistent rise in pneumonia mortality among individuals aged ≥70 years, a gradual increase in ages 50–69, while those under 5 showed a decline. The *n*-sub-epidemic model predicts a decrease in mortality for ages 15–49 and 5–14, contrasting with statistical models showing increasing trends among ages 15–49. The consistency of forecasting trends across models validates the study's findings and indicates their adaptability to different epidemiological contexts.

## Introduction

Pneumonia is a respiratory infection characterized by the build-up of fluids and pus in the alveoli. *Streptococcus pneumoniae* is the most common cause of bacteria associated pneumonia [1]. Certain populations are at a higher risk of developing pneumonia due to underlying factors like age, weakened immune systems, and chronic illnesses [2]. Acknowledging these factors and tackling the underlying causes is essential for the effective management and treatment of pneumonia.

Worldwide estimates from the World Health Organization (WHO) show that pneumonia exerts a profound impact as a leading cause of global mortality across the population. Particularly, among children aged 1 to 5 years, it accounted for about 22% of all deaths in 2019 [3]. Additionally in the same year, adults aged 70 years and older experienced the highest rates of pneumonia globally [4]. This highlights the vulnerability of both young children and the elderly to pneumonia, while mortality rates for other age groups have remained relatively stable over decades [5].

Pneumonia stands as a significant public health concern, causing about million deaths across all age groups [5]. The economic burden of pneumonia is substantial, manifested through increased healthcare expenditures, lost productivity, and strain on the healthcare system [6]. This issue is particularly prominent in middle income countries, which carry a greater burden of pneumonia than higher income countries. As a result, individuals in deprived socioeconomic conditions are at a higher risk of pneumonia [7-10]. The various determinants of pneumonia include education, financial status, age, sex, ethnicity, smoking, and environmental pollutants [11]; thus, making it vital to address pneumonia not just at the clinical front but also at various socioeconomic levels.

The changing epidemiology of lower respiratory infection (LRI) and pneumonia related mortality emphasizes the importance of forecasting pneumonia to predict future morbidity and mortality of the disease. Despite efforts to mitigate at both local and global levels, there still exists a significant burden of disease, especially in developing countries. This necessitates the need for extensive and accurate forecasting of pneumonia mortality to inform effective public health strategies, improve healthcare resource allocation, and enhance policies for pneumonia control. Numerous prediction models have been used in studies to assess their use in clinical settings to mitigate or support in early diagnosis of pneumonia [13-15]. Many different forecasting methodologies, including deep machine learning frameworks [16], time-series [17], and mathematical models [18], have been used in pneumonia modeling. Additionally, infectious disease modeling has expanded in varied disease areas such as COVID-19 [19], Zika [20], mpox [21] and influenza [22]. Regardless of the availability of this multi-faceted forecasting data, limited research addresses predictive analysis of pneumonia mortality across age groups, especially on a global scale.

Our research aims to advance pneumonia mortality forecasting methods by leveraging on models such as Auto-Regressive Integrated Moving Average Models (ARIMA), Generalized Additive Models (GAM), Simple Linear Regression (SLR), Facebook's Prophet Model (Prophet), and the *n*-sub-epidemic modeling framework to forecast global trends in pneumonia mortality across various age groups. It is vital to acknowledge that addressing pneumonia deaths requires tailored public health approaches specific to different age groups, and our study aims to make a substantial contribution in this regard. The forecasts generated from this study can empower policymakers to implement evidence-based interventions and reduce the burden of LRI and preventable pneumonia-related deaths.

## Methods and procedures

### Data

The yearly number of pneumonia deaths data was retrieved from Institute for Health Metrics and Evaluation (IHME), Global Burden of Disease (GBD) 2019 – with minor processing by Our World in Data (downloaded in February 2024). The dataset covered the period from 1990 through 2019, and the primary unit of measure was the number of deaths attributed to pneumonia and other LRIs. The database defined pneumonia based on symptoms alone, thus diseases with similar symptoms could have been included in the database. However, for clarity and simplicity this paper primarily uses the term "pneumonia". Further, the data was cleaned and transposed using Microsoft Excel. Only yearly world data for age groups 70 years and older, 50–69 years, 15–49 years, 5–14 years, and under 5 years were included and data for other regions were excluded. The *StatModPredict*, a R-Studio (v. 4.3.2) dashboard interface [23] was used to generate model forecasts for ARIMA, GAM, SLR and Prophet, while *SubEpiPredict* a MATLAB (v. R2023b) toolbox [24], was utilized to run *n*-sub-epidemic framework.

Approval for this study to be designated as not human subjects research was obtained from Georgia State University, Institutional Review Board.

### Model calibration

For all the models, the calibration period length was set as 30 years, from 1990 through 2019. The forecasting horizon was set at 11-years, which generated yearly forecasts from 2020 to 2030. Additionally, the forecasting period date was set as 2019, which marked the start of the predictive analysis. These settings were kept the same while forecasting for all the different age groups.

### The $n$-sub-epidemic modeling framework

In the *n*-sub-epidemic framework used in this study, each sub-epidemic is modeled by a three-parameter generalized-logistic growth model (GLM). This framework employs multiple epidemic trajectories modeled as the aggregation of overlapping and asynchronous sub-epidemics. A $n$-sub-epidemic trajectory consists of $n$ overlapping sub-epidemics and is described by a system of interconnected differential equations [24]:

$$\frac{dC_i(t)}{dt} = C_i{}'(t) = A_i(t) r_i C_i{}^{p_i}(t)\left(1 - \frac{C_i(t)}{K_{0_i}}\right) \quad (1)$$

The incidence curve of pneumonia deaths is represented by $\frac{dC_i(t)}{dt}$, where $C_i(t)$ is the cumulative number of pneumonia deaths for sub-epidemic $i$. The parameters $r_i, p_i, K_{0_i}$ define the characteristics of the $i_{th}$ sub-epidemic, for $i = 1, \dots, n$. The parameter $r_i$, denotes positive growth rate per unit of time, while parameter $K_0$ signifies the final outbreak size. The early sub-exponential and exponential growth patterns are captured by the "scaling of growth" parameter $p \in [0,1]$. When $p = 0$ , this equation describes a constant number of new deaths over time, while $p = 1$ indicates that the early growth phase is exponential. Intermediate values of $p$ ($0 < p < 1$) indicate early sub-exponential (e.g., polynomial) growth dynamics.

We considered the nonlinear least squares method to estimate the model parameters by matching the model results to the actual observed pneumonia deaths data. Post which, we identified the four best top-ranked sub-epidemic models using the corrected Akaike Information Criterion ($AIC_c$) values from a selection of models with two sub-epidemics. The *n*-sub-epidemic framework quantified parameter uncertainty for the best-fit model, by employing 300 bootstrap realizations approach to project short-term forecasts [24].

For our ensemble models, we created both weighted ensemble (2), (3) and (4) using relative likelihood to assign weights and unweighted ensemble (2), (3) and (4) assuming equal weights of the highest four ranked models to provide different forecasting methodologies, allowing us to compare results and assess the robustness of our projections. The 95% prediction interval (PI) was generated using the bootstrapping approach. More information about the *n*-sub-epidemic framework can be found in reference [24].

**Auto-Regressive Integrated Moving Average Models (ARIMA)**

There are three components involved in ARIMA: the auto-regression (AR) part, which uses the most recent values of the series, the moving average (MA) which accounts for current and past error terms, and the integration (I), which involves differentiation to stabilize the time series by addressing overall trends. The mathematical representation of ARIMA is:

$$\phi(B)(1-B)^d y_t = c + \theta(B)\epsilon_t, \tag{5}$$

Where the number of pneumonia deaths at time $t$ is denoted by $y_t$, backshift operator denoted by $B$, inferred as $By_t = y_{t-1}$ and $B(By_t) = B^2 y_t = y_{t-2}$, etc, $\phi(\mathrm{B}) = 1 - \phi_1 \mathrm{B} - \cdots - \phi_p B^p$ here $p$ is the order of the AR model, $d$ is the degree of differencing, $\theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q$ and the $q$ is the order of the MA model. With this notation, $\phi(\mathrm{B})y_t = y_t - \phi_1 y_{t-1} - \cdots - \phi_p y_{t-p}$, $\theta(B)\epsilon_t = \epsilon_t - \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q}$, and $(1-B)^d$ means conducting the differencing $d$ times. The ARIMA model forecast was run in the *StatModPredict* dashboard using the default settings. Further details on the functions and methods used by the *StatModPredict* dashboard to run ARIMA model forecasts can be found in Bleichrodt A et al., 2024 [23].

**Generalized Additive Models (GAM)**

The generalized additive model (GAM) is a type of Generalized Linear Model including sum of unknown smooth functions of some covariates [25]. Below is the GAM expression specific to our study, where time is the only covariate, assuming that $y_t$ has normal distribution [25]:

$$y_t = \beta_0 + s(t) + \epsilon_t, \tag{6}$$

where $s(.)$ is an unknown smooth function of time, and $\epsilon_t \sim N(0, \sigma^2)$ assuming model normality. To run the GAM model in the *StatModPredict* dashboard we used the default model settings. A more detailed description of the methodology used by the *StatModPredict* dashboard to run GAM model forecasts can be found in Bleichrodt A et al., 2024 [23]

**Facebook's Prophet Model**

Originally created for business forecasting, Facebook's Prophet model [26] has become popular in various disciplines, including infectious disease modeling. The main idea behind the model is that it is "decomposable," meaning that the main trend, $y(t)$, can be broken down into three components (trend [*g(t)*], seasonality [*s(t)*], and holidays or events [*h(t)*]) along with an error term. As a result, the model is structured as follows:

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t \tag{8}$$

In this model, *g*(*t*) represents a non-periodic component that captures the overall time series trend, *s*(*t*) is a periodic trends component that models regular changes over time, and *h*(*t*) accounts for irregular events. The term $\epsilon_t$ represents the model's error at time *t* (normally distributed). We ran the Prophet model in *StatModPredict* dashboard [23] using default settings, including automatic selection of *s*(*t*), described in Taylor S and Letham B et al., 2021 [27].

The SLR model used time as the only predictor, so it is not described in further detail in this paper. An elaborate description of the *StatModPredict* dashboard and the methods associated with fitting and forecasting the statistical models using this interface can be found in Bleichrodt A et al., 2024 [23].

**Forecast interpretation**

**Percent change**

We calculated the percent change to quantify and describe the forecasted change in trends for pneumonia mortality from 2019 through 2030 across different age groups. The calculation was as follows:

$$\frac{Forecasted\ number\ of\ pneumonia\ deaths\ (2030) - Last\ observed\ data\ (2019)}{Last\ observed\ data\ (2019)}\ X\ 100$$

**Model uncertainty**

The model uncertainty in forecasting pneumonia deaths was determined to understand the level of variability in the predicted model trends. This was done by examining the interval trajectory, i.e., the increasing, decreasing or increasing-decreasing trend of the 95% PI bounds. Additionally, the average 95% PI width was calculated by determining the interval width for each year of the forecasting period and then averaging them out.

## Results

### Overall global trend

Figure 1 depicts the global trends in pneumonia mortality per 100000 people since 1990. Notably there has been a consistent and substantial rise in pneumonia-related deaths among individuals aged 70 years and older. Similarly, those aged 50–69 years have also shown a steady increase in mortality over the years. In stark contrast, those below 5 years of age have shown a substantial decline in the number of pneumonia deaths. Meanwhile, the mortality rates for age groups 15–49 years and 5–15 years have remained relatively stable over 30 years.

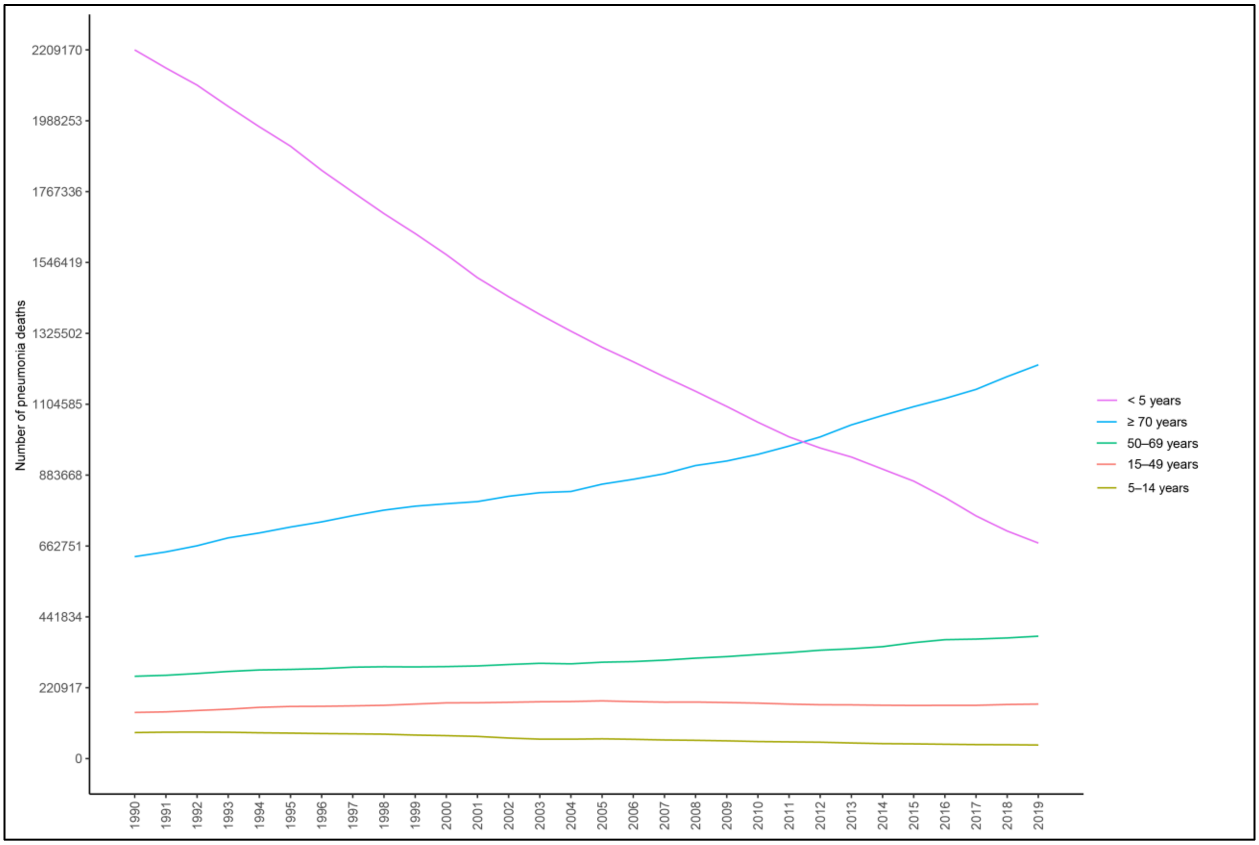


**Figure 1: Global trends in pneumonia mortality per 100000 people across different age groups from 1990 through 2019.** This figure shows the global number of pneumonia deaths among age groups 70 years and older (blue), 50–69 years (green), 15–49 years (red), 5–14 years (yellow), and under 5 years (pink). Each solid-colored line represents the trend in number of pneumonia deaths from 1990 through 2019 for the different age groups. World data for all age groups was obtained from the Institute for Health Metrics and Evaluation (IHME), 2019 [28].

**70 years and older age group**

Overall, our analysis of global trends from 1990 through 2030 indicated an increasing forecast trend in pneumonia mortality for individuals aged 70 years and older, across all models (Figure 2). The ARIMA model produced the widest 95% PI (width: 158238.9) and fanned 95% PI bounds compared to the other statistical models (Figure 2, 8a). Among all the models, the GAM model predicted the highest median deaths (1.6 million [95% PI: 1.5, 1.7]), with a percent change of 35.4% from 2019 to 2030 (Figure 7a). Interestingly, the *n*-sub-epidemic model predicted a lower median number of deaths i.e., 1.2 million (95% PI: 1.1, 1.3) in 2030. This model underpredicted the actual median number of pneumonia-related deaths compared to the observed data in 2019. That said, all the *n*-sub-epidemic top ranked and ensemble models distinctly forecasted an increase in pneumonia mortality for this age group (Figure 2).

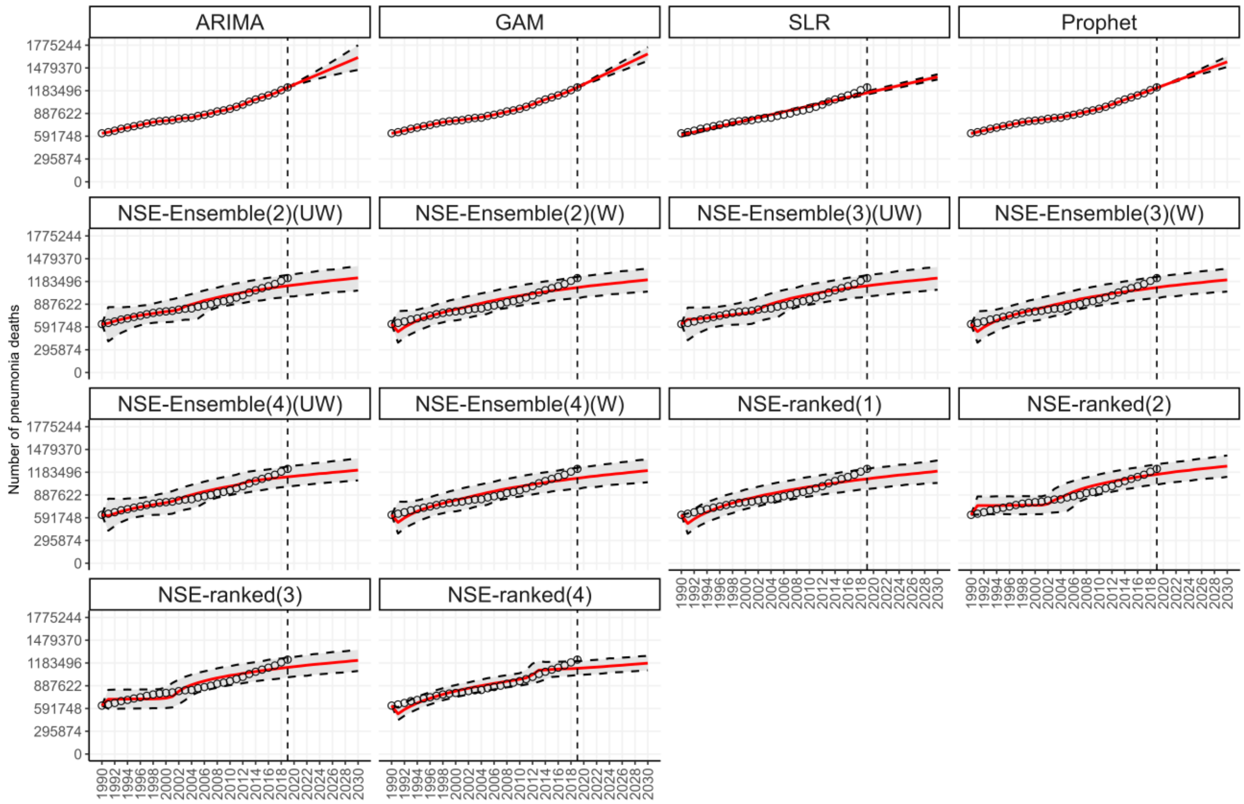


**Figure 2: Model forecasts for pneumonia mortality in the age group 70 years and older, globally.** The panel shows the forecasts generated using statistical models Auto-Regressive Integrated Moving Average Models (ARIMA), Generalized Additive Models (GAM), Simple Linear Regression (SLR), Facebook's Prophet (Prophet) and mathematical *n*-sub-epidemic framework (unweighted (UW), weighted (W) ensemble and top ranked models). The forecasts for all pneumonia-related deaths in the age group 70 years and older, at a global level, are captured. The black circles represent observed data from 1990 to 2019, while the solid red line shows the median forecast. The 95% PI is indicated by the black dashed lines, forming a grey ribbon. The vertical dash line is the forecasted period date, to the right of which represents the start of our predictive analysis.

**50–69 years age group**

Among those aged 50–69 years, similar increasing trends in the number of pneumonia deaths were seen across all models. However, the predicted increase was seen to be gradual over time (Figure 3). Compared to all the models, the Prophet model predicted the highest percent change of 19.9% from 2019 to 2030, and the least percent change of 5.5% was seen with *n*-sub-epidemic third ranked model (Figure 7a, 7b). All models forecasted a consistent increasing trajectory of the upper and lower bound 95% PIs (Figure 8a, 8b). However, the ARIMA and GAM produced slightly wide and fanned 95% PI, estimating some forecasting uncertainty until 2030. All the *n*-sub-epidemic top ranked and ensemble models forecasted a consistent increasing trend within narrow 95% PI bounds.

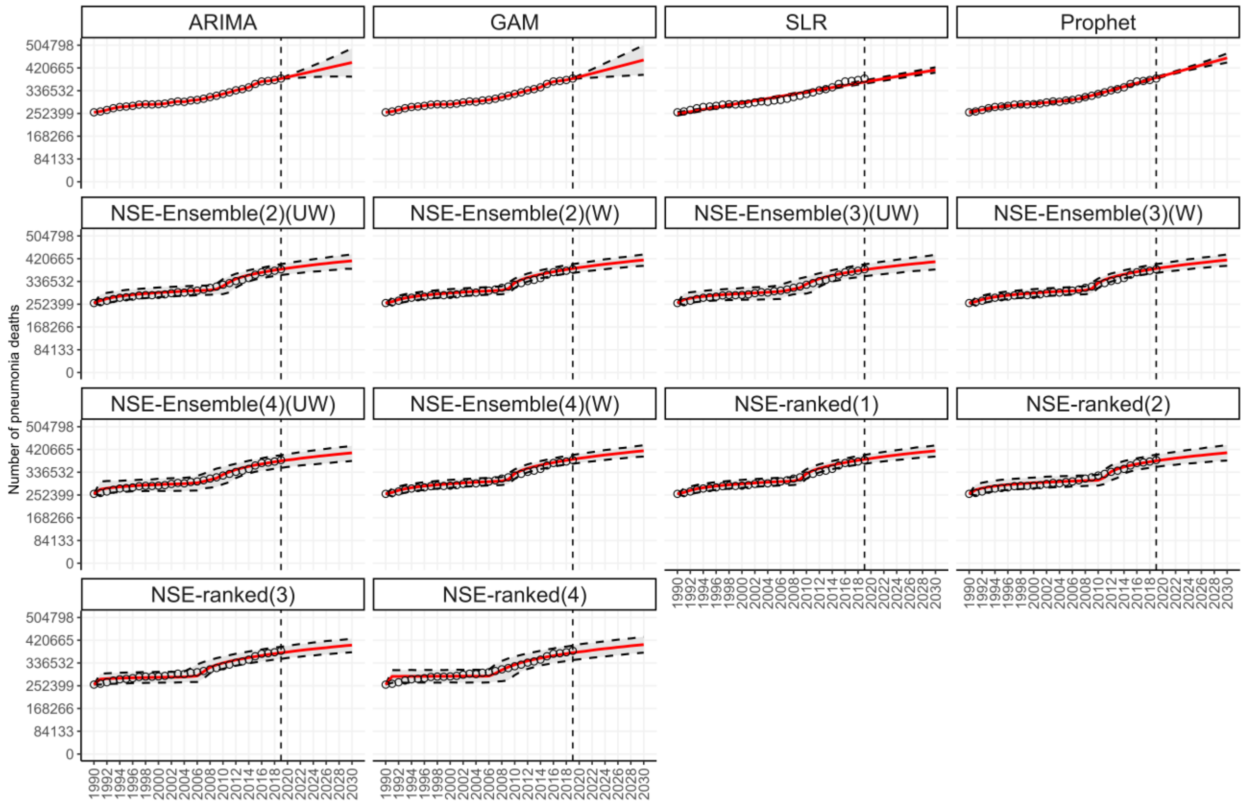


**Figure 3: Model forecasts for pneumonia mortality in the age group 50–69 years, globally.** The panel shows the forecasts generated using statistical models Auto-Regressive Integrated Moving Average Models (ARIMA), Generalized Additive Models (GAM), Simple Linear Regression (SLR), Facebook's Prophet (Prophet) and mathematical *n*-sub-epidemic framework (unweighted (UW), weighted (W) ensemble and top ranked models). The forecasts for all pneumonia-related deaths in the age group 50–69 years, at a global level, are captured. The black circles represent observed data from 1990 to 2019, while the solid red line shows the median forecast. The 95% PI is indicated by the black dashed lines, forming a grey ribbon. The vertical dash line is the forecasted period date, to the right of which represents the start of our predictive analysis.

### 15 – 49 years age group

For the age group 15–49 years, figure 4 showed some variations among model forecasts. All four statistical models indicated an increasing percent change ranging from 0.1% to 11.6% (Figure 7a), whereas the *n*-sub-epidemic models predict a decreasing percent change, ranging from −9.1% to −11.6% (Figure 7b), with narrow 95% PIs (Figure 4). While the statistical models display an upward trend, their 95% prediction intervals were fanned (i.e. increasing-decreasing interval trajectory), leading to heightened uncertainty in the related forecasts. Additionally, the ARIMA model had the largest 95% PI width of 44555.4 (Figure 8a), while the *n*-sub-epidemic second ranked model had the smallest width at 13670.6 (Figure 8b).

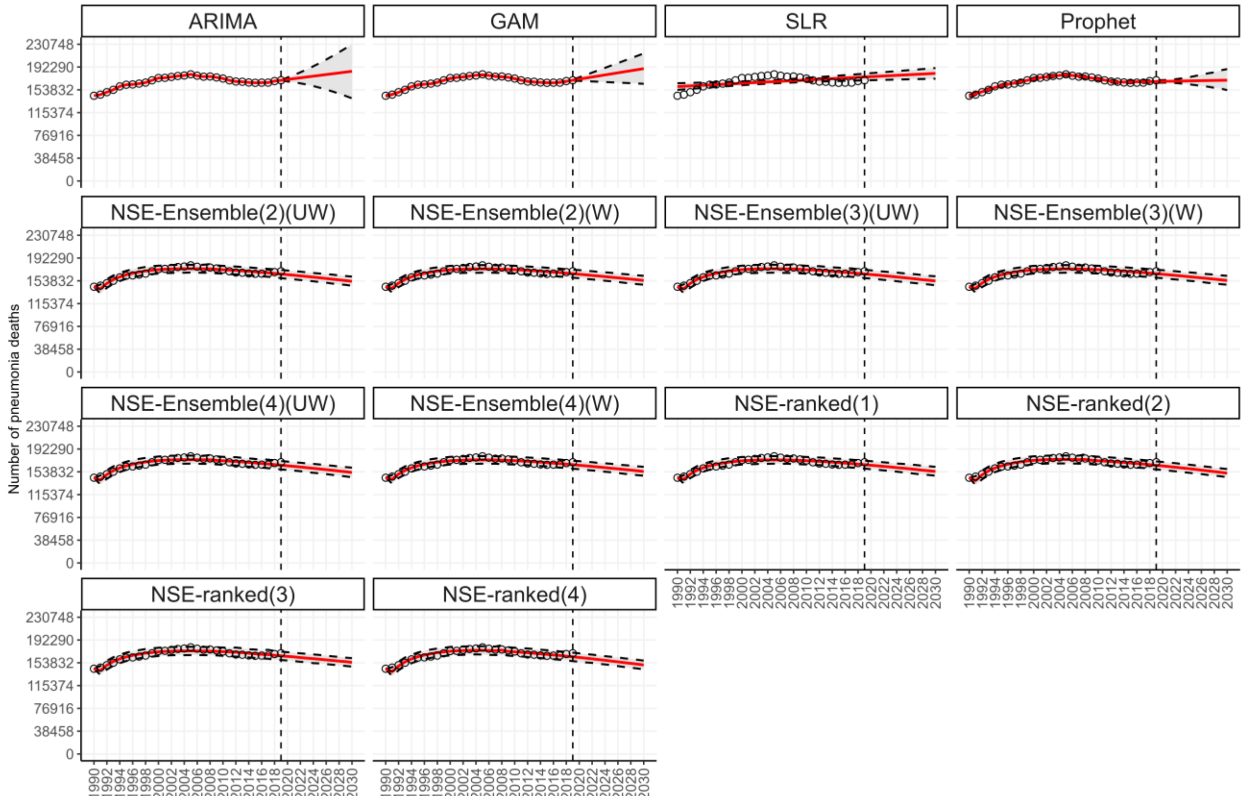


**Figure 4: Model forecasts for pneumonia mortality in the age group 15–49 years, globally.** The panel shows the forecasts generated using statistical models Auto-Regressive Integrated Moving Average Models (ARIMA), Generalized Additive Models (GAM), Simple Linear Regression (SLR), Facebook's Prophet (Prophet) and mathematical *n*-sub-epidemic framework (unweighted (UW), weighted (W) ensemble and top ranked models). The forecasts for all pneumonia-related deaths in the age group 15–49 years, at a global level, are captured. The black circles represent observed data from 1990 to 2019, while the solid red line shows the median forecast. The 95% PI is indicated by the black dashed lines, forming a grey ribbon. The vertical dash line is the forecasted period date, to the right of which represents the start of our predictive analysis.

**5–14 years age group**

All models predicted a consistent decline in pneumonia mortality trends for individuals aged 5–14 years (Figure 5). The statistical models showed a percent change ranging from −25.7% to −45.9% (Figure 7a), while the *n*-sub-epidemic models exhibited a percent change from −22.7% to −29.2% (Figure 7b). The widest forecast, 95% PIs were observed with the ARIMA (14983.1) and GAM (27447.7) models, where the GAM model displayed some level of forecasting uncertainty in the 95% PI trajectory (i.e. increasing-decreasing interval trajectory) (Figure 8a). On the other hand, among the *n*-sub-epidemic models the second ranked model had a relatively narrow 95% PI width of 5913.2 (Figure 8b).

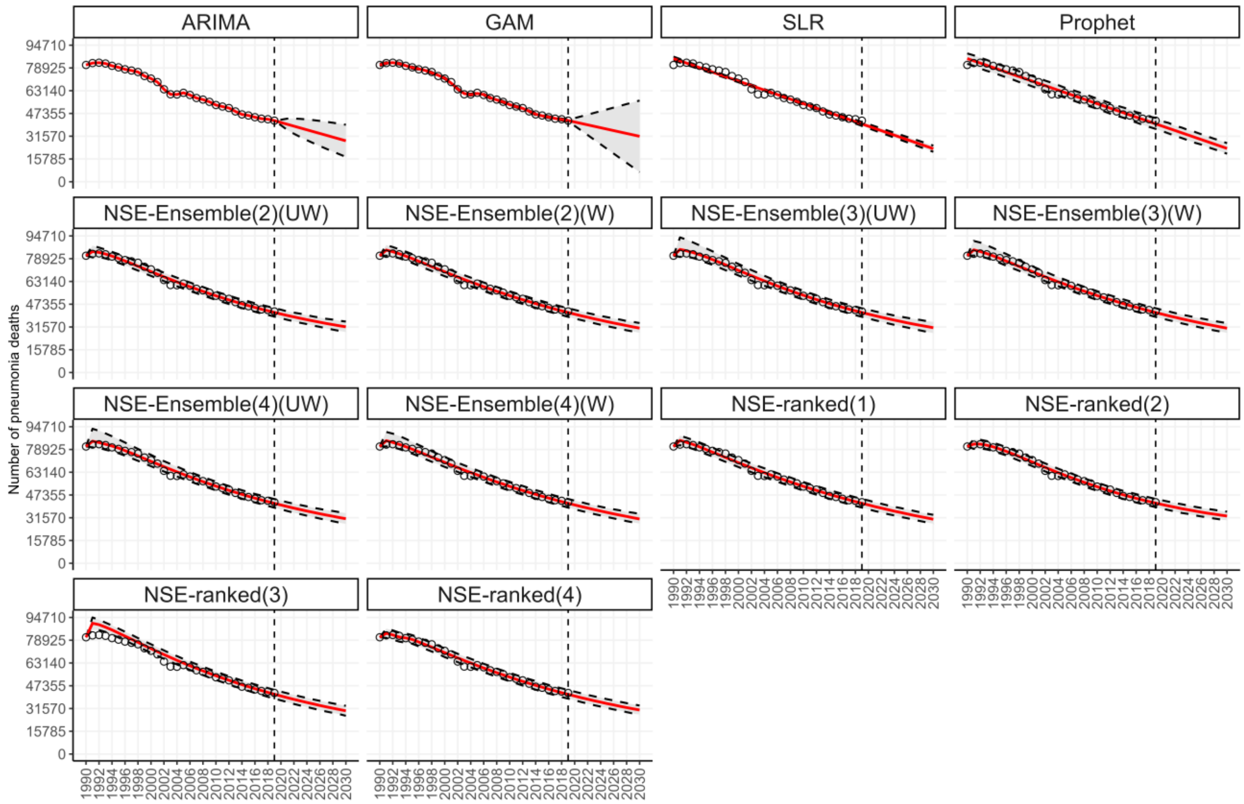


**Figure 5: Model forecasts for pneumonia mortality in the age group 5–14 years, globally.** The panel shows the forecasts generated using statistical models Auto-Regressive Integrated Moving Average Models (ARIMA), Generalized Additive Models (GAM), Simple Linear Regression (SLR), Facebook's Prophet (Prophet) and mathematical *n*-sub-epidemic framework (unweighted (UW), weighted (W) ensemble and top ranked models). The forecasts for all pneumonia related deaths in the age group 5–14 years, at a global level, are captured. The black circles represent observed data from 1990 to 2019, while the solid red line shows the median forecast. The 95% PI is indicated by the black dashed lines, forming a grey ribbon. The vertical dash line is the forecasted period date, to the right of which represents the start of our predictive analysis.

**Under 5 years age group**

Forecast for those under 5 years of age showed a declining trend across all statistical and *n*-sub-epidemic models, with very narrow 95% PIs (Figure 6). The overall percent change in forecasted deaths by the statistical models ranged from −57.2% to −97.3%, with the SLR model predicting the highest percent change from 2019 to 2030 (Figure 7a). The *n*-sub-epidemic models on the other hand showed consistent percent change between −33.2% to −33.9% (Figure 7b). The widest forecast interval width of 289912.0 was observed with the ARIMA (Figure 8a), while the *n*-sub-epidemic first ranked model showed the smallest interval width of 58729.0 compared to all the other models (Figure 8b).

In this study the ARIMA model captured the linear or nonlinear (specifically, quadratic) trends by choosing the differencing order d=1 or 2 and the AR order p=0, indicating that after appropriate differencing, the previous observations of the time series do not provide additional information for pneumonia mortality.

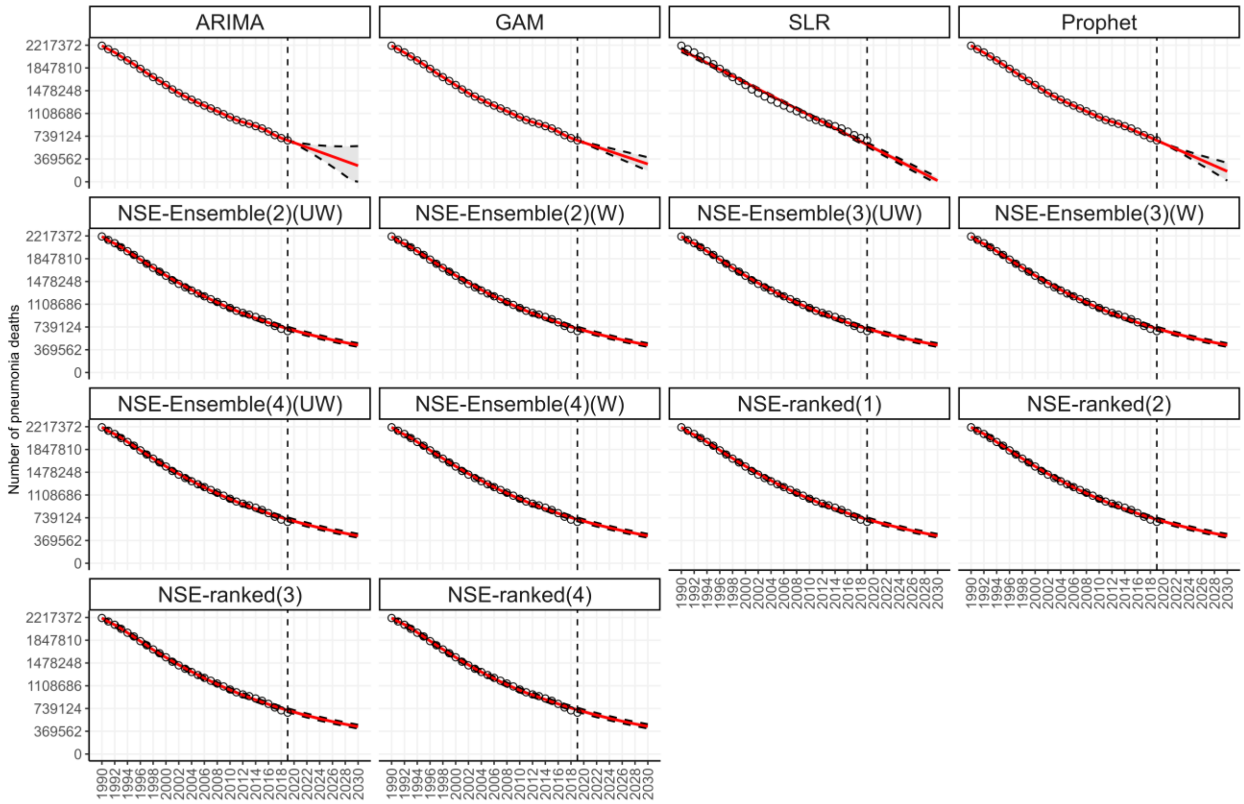


**Figure 6: Model forecasts for pneumonia mortality in the age group under 5 years, globally.** The panel shows the forecasts generated using statistical models Auto-Regressive Integrated Moving Average Models (ARIMA), Generalized Additive Models (GAM), Simple Linear Regression (SLR), Facebook's Prophet (Prophet) and mathematical *n*-sub-epidemic framework (unweighted (UW), weighted (W) ensemble and top ranked models). The forecasts for all pneumonia related deaths in the age group under 5 years, at a global level, are captured. The black circles represent observed data from 1990 to 2019, while the solid red line shows the median forecast. The 95% PI is indicated by the black dashed lines, forming a grey ribbon. The vertical dash line is the forecasted period date, to the right of which represents the start of our predictive analysis.

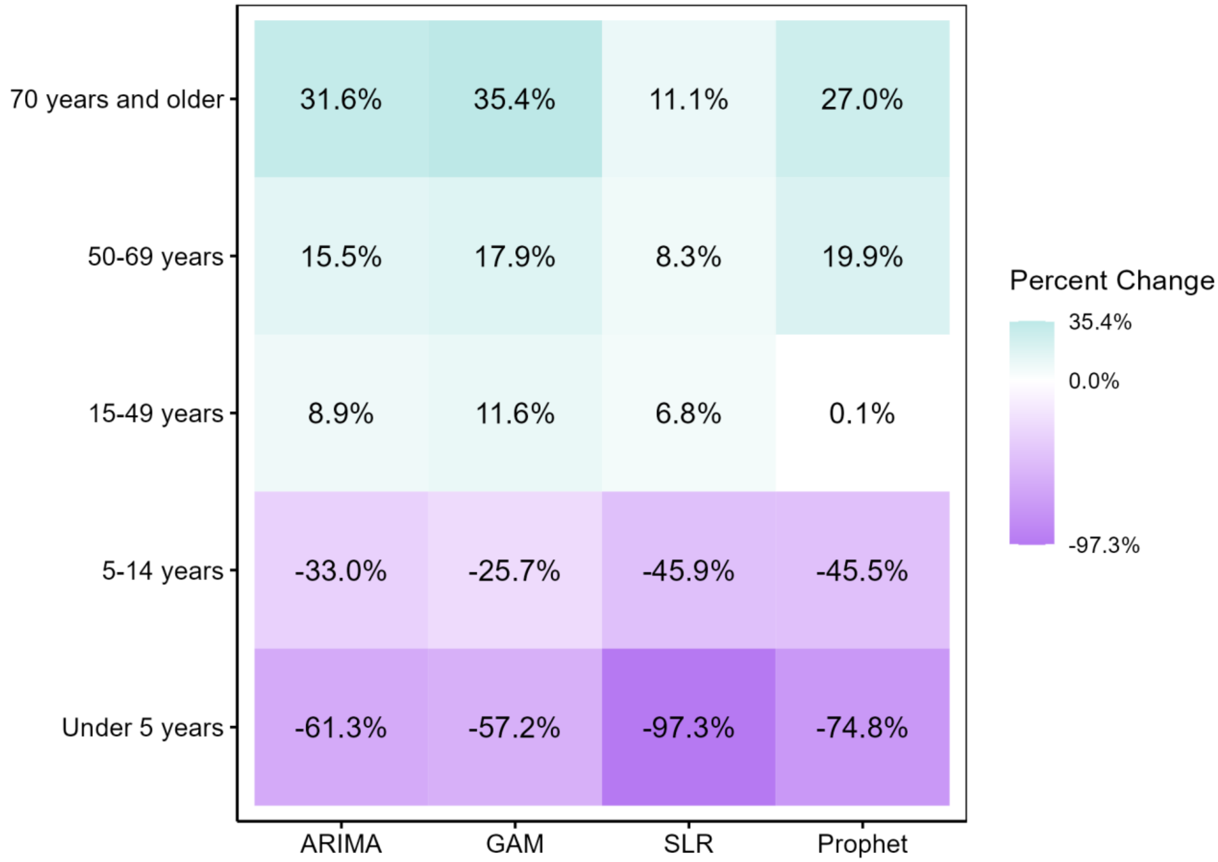


**Figure 7 (a): Percent change in the number of pneumonia deaths from 2019 to 2030 for age groups 70 years and older, 50–69 years, 15–49 years, 5–14 years, and under 5 years, generated using statistical models.** The panel shows the relative percent change from the last year of data used during the model calibration process (2019) to the forecasted median number of pneumonia deaths in 2030 for different age groups. This figure shows results from the Auto-Regressive Integrated Moving Average Models (ARIMA), Generalized Additive Models (GAM), Simple Linear Regression (SLR), and Facebook's Prophet model. Purple tiles indicate forecasted decreases, and green tiles indicate forecasted increases from 2019 to 2030. Darker colors indicate a greater percent change in either direction (i.e., negative or positive). The values within each tile correspond to the relative percent change in pneumonia deaths from 2019 to 2030.

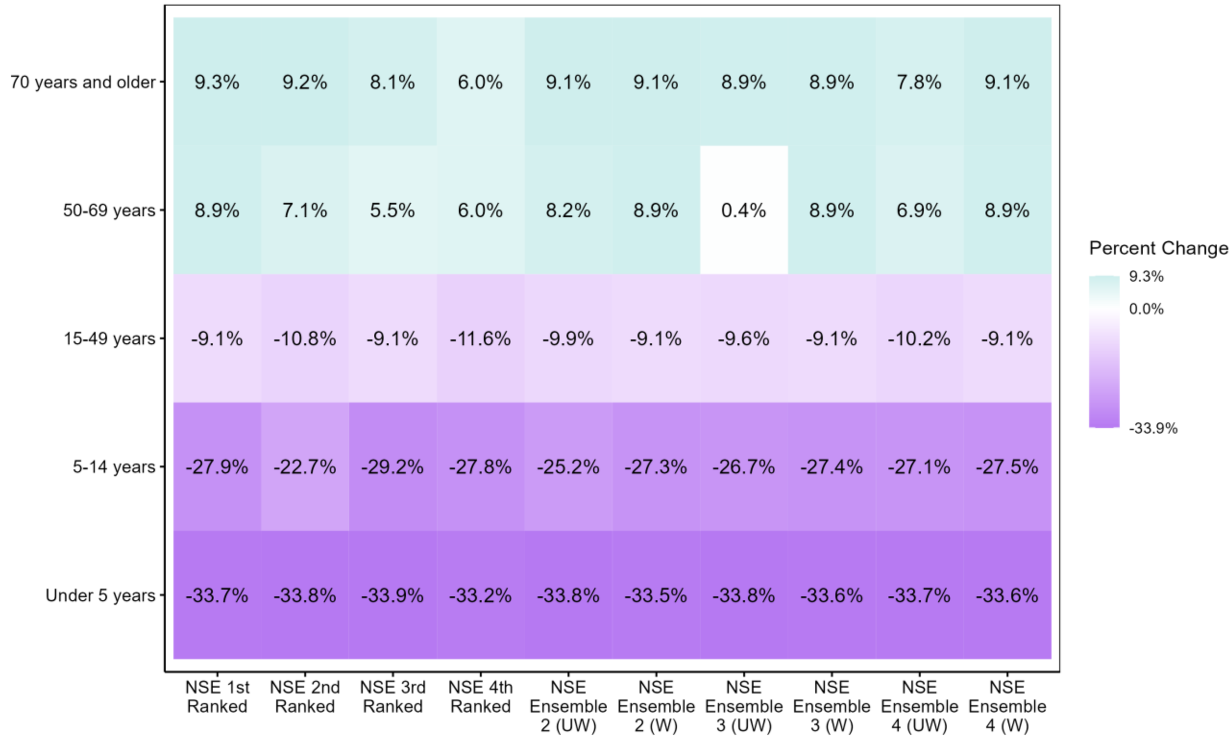


**Figure 7 (b): Percent change in the number of pneumonia deaths from 2019 to 2030 for age groups 70 years and older, 50–69 years, 15–49 years, 5–14 years, and under 5 years, generated using mathematical modeling framework.** The panel shows the relative percent change from the last year of data used during the model calibration process (2019) to the forecasted median number of pneumonia deaths in 2030 for different age groups. This figure shows results from the *n*-sub-epidemic (NSE) top ranked, unweighted (UW), and weighted (W) ensemble models. Purple tiles indicate forecasted decreases, and green tiles indicate forecasted increases from 2019 to 2030. Darker colors indicate a greater percent change in either direction (i.e., negative or positive). The values within each tile correspond to the relative percent change in pneumonia deaths from 2019 to 2030.

*The Percent change for age group 70 years and older was calculated against median model fit 2019 instead of observed data.*

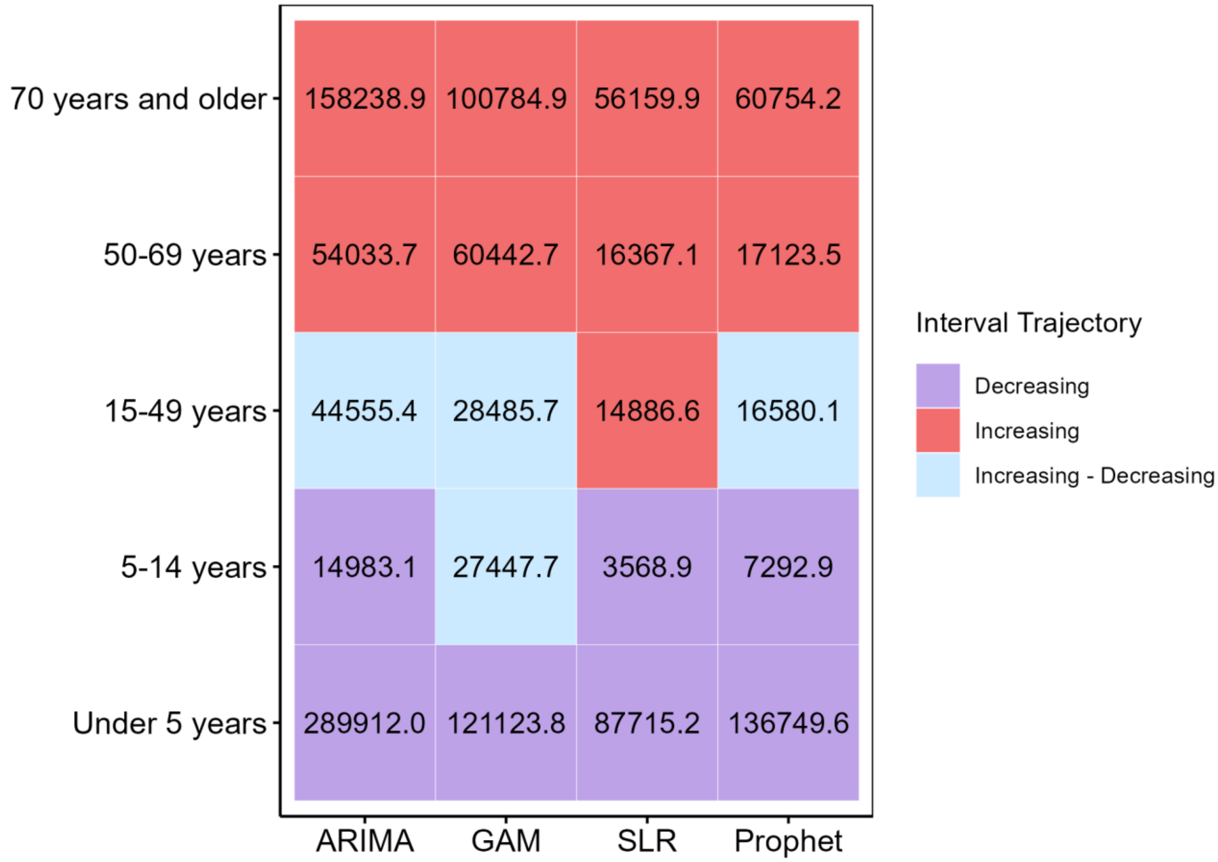


**Figure 8 (a): Statistical model uncertainty associated with pneumonia mortality forecasts among age groups 70 years and older, 50–69 years, 15–49 years, 5–14 years, and under 5 years.** The panel shows the uncertainty related to the number of pneumonia deaths forecasts (2020 to 2030) associated with the statistical models Auto-Regressive Integrated Moving Average Models (ARIMA), Generalized Additive Models (GAM), Simple Linear Regression (SLR), and Facebook's Prophet model. Data is provided by age groups 70 years and older, 50–69 years, 15–49 years, 5–14 years, and under 5 years. The colors shown correspond to the directions of the upper and lower 95% PIs for each forecast. "Decrease" indicates that both bounds decrease from the start to end of the forecasting period, "Increase" indicates both bounds increase from start to end, and "Increase-Decrease" describes contradicting trends between the upper and lower bounds. The values shown within each tile represent the average 95% PI width across the 11-year forecast period.

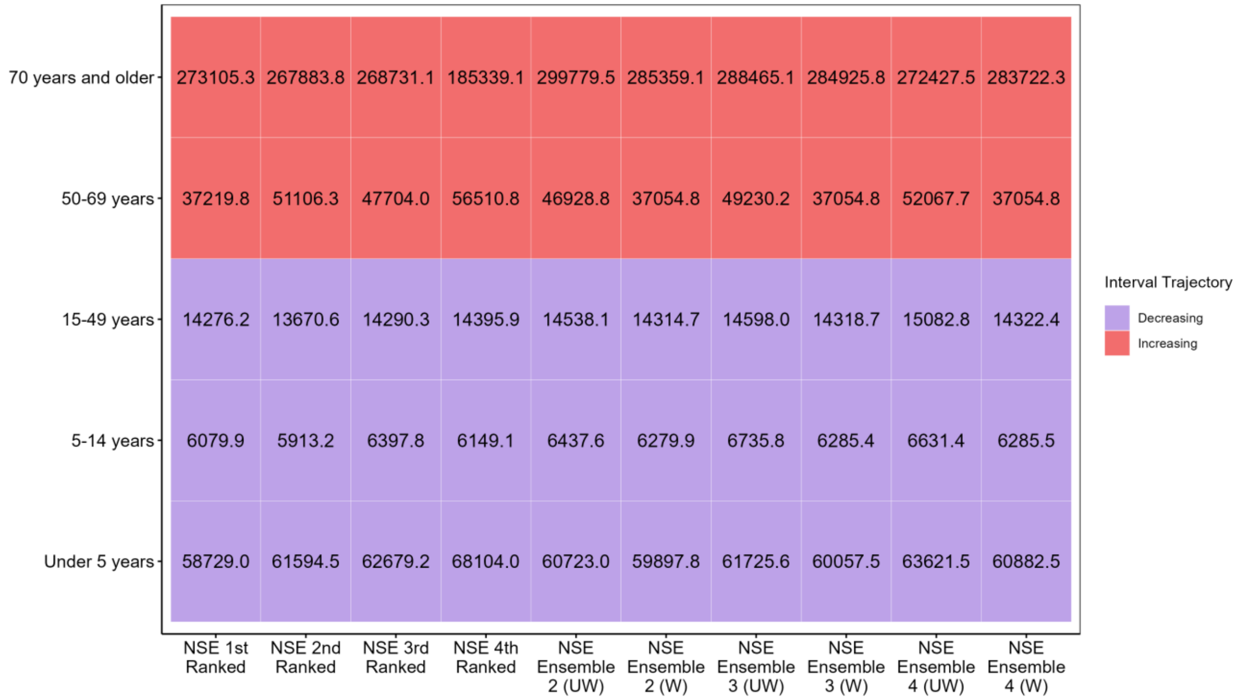


**Figure 8 (b): Mathematical model uncertainty associated with pneumonia mortality forecasts among age groups 70 years and older, 50–69 years, 15–49 years, 5–14 years, and under 5 years.** The panel shows the uncertainty related to the number of pneumonia deaths forecasts (2020 to 2030) generated using the mathematical *n*-sub-epidemic framework (top ranked, unweighted (UW), and weighted (W) ensemble models). Data is provided by age groups 70 years and older, 50–69 years, 15–49 years, 5–14 years, and under 5 years. The colors shown correspond to the directions of the upper and lower 95% PIs for each forecast. "Decrease" indicates that both bounds decrease from the start to end of the forecasting period, "Increase" indicates both bounds increase from start to end, and "Increase-Decrease" describes contradicting trends between the upper and lower bounds. The values shown within each tile represent the average 95% PI width across the 11-year forecast period.

## Discussion

The findings from our research shed light on forecast trends in pneumonia mortality across different age groups, using ARIMA, GAM, SLR, Prophet and *n*-sub-epidemic forecasting models. These models were selected for their distinct strengths in handling time-series data and their complementary abilities to capture different aspects of pneumonia mortality trends. Specifically, the ARIMA models capture linear or nonlinear trends. The GAM and Prophet models provide flexible methodologies to capture nonlinear relationships. Finally, the *n*-sub-epidemic framework models overlapping and asynchronous sub-epidemic trajectories, making it suitable for heterogeneous patterns in infectious disease mortality trends. By including both statistical (ARIMA, GAM, Prophet) and mechanistic (*n*-sub-epidemic) approaches, we ensured a comprehensive analysis that leverages on model strengths to validate trends and improve forecast reliability.

The GBD 2021 estimated that among those aged 70 years and over the incidence rate of non-COVID-19 LRIs globally was the highest [29]. In support of which, our model forecasts found that the number of pneumonia deaths are predicted to escalate from around 1 million in 2019 to about 1.6 million by 2030, among this age group. Similar increasing trends were reported among the 50-69 years age group in our study. One contributing factor to this concerning surge in pneumonia mortality could be the projected demographic shift towards an aging population [5,30-31]. Along with aging, diminished immunity, deteriorating lung capacity, coexistence of comorbidities [31], and exposure to environmental particulate matter [32] have been common contributing factors for pneumonia related mortality among the elderly. Some studies have also found positive association between climatic changes like humidity, temperature variation, extreme cold, temperate climate and incidence of pneumonia [33-34].

This grave burden of pneumonia is majorly recorded in low- and middle-income countries, with rates higher than the global average of 60 deaths per 100000 among children [29]. Additional estimates by Shi T et al., in their study reported about 600000 more pneumonia related hospitalizations among the elderly in developing countries, compared to industrialized countries [35]. The low- and middle-income countries often experience higher mortality rates, exacerbated by poor vaccine access, limited healthcare facilities, and greater exposure to environmental pollutants. Conversely, high-income countries benefit from advanced healthcare systems and widespread immunization programs, resulting in lower mortality rates.

In our study all statistical and the *n*-sub-epidemic models forecast that those under 5 years of age will experience the highest declining percent change in pneumonia mortality by 2030, highlighting the positive outcomes of policies and interventions focusing on this age group. This is in accordance with continued steady declining trend in pneumonia deaths from 1990 to 2019, among this age group, with about a 28% decline in the number of deaths specifically between 2019 to 2021 [29]. This drastic drop was mainly due to the restrictions and preventive measures incorporated to lower COVID-19. The implemented strategies, such as hand-hygiene, isolation, lockdowns and facemasks proved beneficial in reducing the overall incidence of respiratory infections across all ages [29]. Moreover, the IHME 2021 also projected that LRIs will become the 9th leading cause of global disease burden by 2050, giving up its current 5th position [36]. This reduction could be attributed to many factors such as the reduction in childhood malnutrition, improved air pollution policies, better sanitation, smoke-free laws, overall improved access to health care, health technology, medication and vaccines [5,37-38].

Furthermore, emergence of vaccines as an effort to prevent pneumonia and other related respiratory infections has shown favorable reduction in deaths due to pneumonia [29,39]. However, certain disparities in vaccine coverage do exist especially in low- and middle- income countries. A model-based epidemiological study suggested that efforts in introduction of pneumococcal conjugate vaccine (PCV) in low- and middle – income countries could prevent about 697000 deaths, highlighting the importance of improving vaccine coverage in developing countries [40].

Antimicrobial resistance is another significant threat to global health, killing about 4.95 million each year [41]. Resistance to antibiotics is a serious problem that makes the treatment of pneumonia harder, resulting in longer hospitalization, and complicating patient care. A study conducted in Bangladesh among children under 5 years of age observed that children who were infected with antimicrobial-resistant bacteria had higher rate of mortality [42]. The elderly population are also highly susceptible to drug-related pneumonia mortality due to existing comorbidities and weaker immune system [43]. This rising resistance to common antibiotics leads to higher treatment failure and adverse pneumonia outcomes. The emergence of vaccine and employment of antibiotic stewardship programs could help reduce pneumonia related deaths over the years as a solution to overcome these challenges with antimicrobial resistance.

The incidence and mortality rate of pneumonia among 15–49 years and children 5–14 years, has remained consistent. However, limited studies focused on understanding the trends in pneumonia mortality in these age groups [4]. All our model forecast predicted that by 2030 there will be a consistent decrease in pneumonia mortality, among 5–14 years age group, globally. Although, certain uncertainties in model forecasts were noted in forecasting pneumonia mortality. For example, forecasts for the 15–49 age group showed conflicting trends, with statistical models (ARIMA, GAM, Prophet) predicting slight increase in pneumonia mortality, while *n*-sub-epidemic models forecasted a decrease. This discrepancy may stem from inherent differences in the models' assumptions and methodologies. Statistical models like ARIMA and GAM rely on historical trends and may extrapolate small upward trends, potentially overestimating future mortality. In contrast, the *n*-sub-epidemic framework accounts for overlapping epidemic dynamics, possibly capturing declining trends in specific sub-epidemics not apparent in the aggregate data.

Several factors could drive these divergent predictions. Improvements in healthcare access, vaccination coverage, and socioeconomic conditions could contribute to declining trends, as captured by the *n*-sub-epidemic models, while statistical models may reflect persistent vulnerabilities such as gaps in healthcare or antimicrobial resistance that disproportionately affect this working-age population. Additionally, environmental and behavioral factors like urbanization and smoking prevalence might exert varying influences on mortality trends across regions and populations within this age group.

To better understand these discrepancies, future work could incorporate additional data sources, such as region-specific socioeconomic and healthcare indicators, to refine model predictions. Exploring interactions between key drivers like vaccination uptake and antimicrobial resistance could provide a more nuanced understanding of the factors influencing mortality trends in the 15–49 age group. Addressing these complexities could yield more accurate and actionable insights for targeted interventions.

This study presents certain limitations. The database definition of pneumonia was derived based on symptoms alone, thus other LRI with similar symptoms may have been included. Additionally, our study focused on global pneumonia mortality trends, and we acknowledge that regional variations are crucial to understanding the full scope of the disease burden. Pneumonia mortality rates can vary significantly across regions due to disparities in socioeconomic conditions, healthcare access, environmental factors, and public health infrastructure. Future studies could extend this work by incorporating region-specific data to explore these variations in greater detail. This would allow for tailored public health interventions addressing the unique challenges and resources of different regions. It should also be noted that the statistical models were run using default *StatModPredict* dashboard settings without any customizations for growth trends, seasonal changes, and events or holidays. Moreover, the *n*-sub-epidemic framework did not account for behavioral changes or interventions which could impact disease trajectory. Further analysis is required to assess the variation in trends after accounting for these factors. Retrospective analysis of this study should be considered in the future once data for the forecasted years are available to determine the robustness and accuracy of these forecasts. The reliability of our model forecasts was solely based on the consistency across all the models in predicting the number of pneumonia deaths in different age groups and observing the trends based on historic data. It is important to note that as this study did not evaluate the forecast model performance metrics, thus a definite conclusion on which model performed the best cannot be drawn.

Despite these limitations, the global trends identified in our study offer a foundational understanding of pneumonia mortality upon which future research can build. The insights gained from our research can offer valuable assistance for current and future public health planning, as the predicted trends reflect and highlight the efficacy of current intervention strategies and their effect on future outcomes. Besides, data generated from this study could be utilized to shift focus towards improving vaccination coverage, treatment and other interventions towards those age groups that are predicted to see higher pneumonia mortality over the coming years.

## Conclusion

Our study indicated that by 2030 the number of pneumonia deaths among the elderly is predicted to increase, while in children below 5 years it will decrease. This understanding highlights the importance of channeling efforts towards age-specific interventions to effectively mitigate pneumonia mortality. Additionally, this study also highlighted the importance of robust modeling techniques in long-term pneumonia mortality forecasting at a global level under varying assumptions and data characteristics. Future research should consider incorporating additional data sources, refining model parameters, and evaluating model performance metrics to enhance predictive accuracy of the models and utility of results in informing public health strategies.

## References


1. **National Institute of Health** (2022) What Is Pneumonia? Available at https://www.nhlbi.nih.gov/health/pneumonia (accessed January, 14 2024).
2. **Centers for Disease Control and Prevention** (2022) Pneumonia. Available at https://www.cdc.gov/pneumonia/index.html (accessed January 14 2024).
3. **World Health Organization** (2022) Pneumonia in children. Available at https://www.who.int/news-room/fact-sheets/detail/pneumonia (accessed January 15 2024).
4. **GBD Lower Respiratory Infections Collaborators** (2022) Age-sex differences in the global burden of lower respiratory infections and risk factors, 1990-2019: results from the Global Burden of Disease Study 2019. *Lancet Infect Dis*; **22**: 1626-1647. https://doi.org/10.1016/S1473-3099(22)00510-2.
5. **Dadonaite B. and Roser M.** (2019) Pneumonia. Available at https://ourworldindata.org/pneumonia (accessed January 15 2024).
6. **Broulette J, et al.** (2013) The incidence rate and economic burden of community-acquired pneumonia in a working-age population. *Am Health Drug Benefits*; **6**: 494-503.
7. **Azab SF, et al.** (2014) Impact of the socioeconomic status on the severity and outcome of community-acquired pneumonia among Egyptian children: a cohort study. *Infect Dis Poverty*; **3**: 14. https://doi.org/10.1186/2049-9957-3-14.
8. **Calvillo-King L, et al.** (2013) Impact of social factors on risk of readmission or mortality in pneumonia and heart failure: systematic review. *J Gen Intern Med*; **28**: 269-282. https://doi.org/10.1007/s11606-012-2235-x.
9. **Cashat-Cruz M, Morales-Aguirre JJ, Mendoza-Azpiri M** (2005) Respiratory tract infections in children in developing countries. *Semin Pediatr Infect Dis*; **16**: 84-92. https://doi.org/10.1053/j.spid.2005.12.005.
10. **Thorn LK, et al.** (2011) Pneumonia and poverty: a prospective population-based study among children in Brazil. *BMC Infect Dis*; **11**: 180. https://doi.org/10.1186/1471-2334-11-180.
11. **Saunders M, Nellums L** (2022) The indirect effects of COVID-19 upon global childhood pneumonia. *Public Health Pract (Oxf)*; **3**: 100228. https://doi.org/10.1016/j.puhip.2022.100228.
12. **Every Breath Counts** Every Breath Counts Coalition. Available at https://stoppneumonia.org/ (accessed January 14 2024).
13. **Effah CY, et al.** (2022) Machine learning-assisted prediction of pneumonia based on non-invasive measures. *Front Public Health*; **10**: 938801. https://doi.org/10.3389/fpubh.2022.938801.
14. **Liz H, et al.** (2021) Ensembles of Convolutional Neural Network models for pediatric pneumonia diagnosis. *Future Generation Computer Systems*; **122**: 220-233. https://doi.org/10.1016/j.future.2021.04.007.
15. **Schierenberg A, et al.** (2016) External Validation of Prediction Models for Pneumonia in Primary Care Patients with Lower Respiratory Tract Infection: An Individual Patient Data Meta-Analysis. *PLoS One*; **11**: e0149895. https://doi.org/10.1371/journal.pone.0149895.
16. **Ayan E, Karabulut B, Unver HM** (2022) Diagnosis of Pediatric Pneumonia with Ensemble of Deep Convolutional Neural Networks in Chest X-Ray Images. *Arab J Sci Eng*; **47**: 2123-2139. https://doi.org/10.1007/s13369-021-06127-z.
17. **Hussein HA, Mahmoud MM, Eisa HA** (2023) Performance Evaluation of ARIMA and FB-Prophet Forecasting Methods in the Context of Endemic Diseases: A Case Study of Gedaref State in Sudan. *EAI Endorsed Transactions on Smart Cities*; **7**: e1-e1.

18. **Ni Kadek Ary Indah S, Oka S, Ayu W** (2022) Forecasting Pneumonia Toddler Mortality Using Comparative Model ARIMA and Multilayer Perceptron. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*; **6**: 528-537. https://doi.org/10.29207/resti.v6i4.4106.
19. **Chowell G, et al.** (2022) An ensemble n-sub-epidemic modeling framework for short-term forecasting epidemic trajectories: Application to the COVID-19 pandemic in the USA. *PLoS Comput Biol*; **18**: e1010602. https://doi.org/10.1371/journal.pcbi.1010602.
20. **Oidtman RJ, et al.** (2021) Trade-offs between individual and ensemble forecasts of an emerging infectious disease. *Nature communications*; **12**: 5379.
21. **Bleichrodt A, et al.** (2024) Evaluating the forecasting performance of ensemble sub-epidemic frameworks and other time series models for the 2022–2023 mpox epidemic. *Royal Society Open Science*; **11**. https://doi.org/10.1098/rsos.240248.
22. **Reich NG, et al.** (2019) Accuracy of real-time multi-model ensemble forecasts for seasonal influenza in the US. *PLoS computational biology*; **15**: e1007486.
23. **Bleichrodt A, et al.** (2024) StatModPredict: A User-Friendly R-Shiny Interface for Fitting and Forecasting with Statistical Models. *Available at SSRN 4849702*.
24. **Chowell G, et al.** (2024) SubEpiPredict: A tutorial-based primer and toolbox for fitting and forecasting growth trajectories using the ensemble n-sub-epidemic modeling framework. *Infect Dis Model*; **9**: 411-436. https://doi.org/10.1016/j.idm.2024.02.001.
25. **Wood SN.** (n.d.) GAM: Generalized additive models with integrated smoothness estimation. Available at https://www.rdocumentation.org/packages/mgcv/versions/1.9-1/topics/gam. (accessed).
26. **Battineni G, Chintalapudi N, Amenta F** (2020) Forecasting of COVID-19 epidemic size in four high hitting nations (USA, Brazil, India and Russia) by Fb-Prophet machine learning model. *Applied Computing and Informatics*.
27. **Taylor S. & Letham B.** (2021) Package ‘prophet’. Available at https://cran.r-project.org/web/packages/prophet/prophet.pdf (accessed).
28. **Institute for Health Metrics and Evaluation (IHME) GBoDS** (2019) Global Burden of Disease (GBD). In.: IHME, Global Burden of Disease Study.
29. **GBD LRIs & Antimicrobial Resistance Collaborators** (2024) Global, regional, and national incidence and mortality burden of non-COVID-19 lower respiratory infections and aetiologies, 1990-2021: a systematic analysis from the Global Burden of Disease Study 2021. *Lancet Infect Dis*; **24**: 974-1002. https://doi.org/10.1016/S1473-3099(24)00176-2.
30. **World Health Organization: WHO.** (2022) Ageing and health. Available at https://www.who.int/news-room/fact-sheets/detail/ageing-and-health (accessed April 10 2024).
31. **Zhang ZX, et al.** (2018) Prognostic factors for mortality due to pneumonia among adults from different age groups in Singapore and mortality predictions based on PSI and CURB-65. *Singapore Med J*; **59**: 190-198. https://doi.org/10.11622/smedj.2017079.
32. **Croft DP, et al.** (2019) The association between respiratory infection and air pollution in the setting of air quality policy and economic change. *Annals of the American Thoracic Society*; **16**: 321-330.
33. **Makrufardi F, et al.** (2024) Extreme temperatures increase the risk of pediatric pneumonia: a systematic review and meta-analysis. *Frontiers in Pediatrics*; **12**: 1329918.
34. **Miyayo SF, et al.** (2021) Analysis of pneumonia occurrence in relation to climate change in Tanga, Tanzania. *International Journal of Environmental Research and Public Health*; **18**: 4731.
35. **Shi T, et al.** (2020) Global and Regional Burden of Hospital Admissions for Pneumonia in Older Adults: A Systematic Review and Meta-Analysis. *J Infect Dis*; **222**: S570-S576. https://doi.org/10.1093/infdis/jiz053.

36. **Institute for Health Metrics and Evaluation (IHME)** (2021) Global Burden of Disease 2021: Findings from the GBD 2021 Study., 2024. Seattle, WA. Available at https://www.healthdata.org/research-analysis/library/global-burden-disease-2021-findings-gbd-2021-study (accessed.
37. **Dowell SF, et al.** (2000) Mortality from pneumonia in children in the United States, 1939 through 1996. *N Engl J Med*; **342**: 1399-1407. https://doi.org/10.1056/NEJM200005113421904.
38. **McAllister DA, et al.** (2019) Global, regional, and national estimates of pneumonia morbidity and mortality in children younger than 5 years between 2000 and 2015: a systematic analysis. *The Lancet Global Health*; **7**: e47-e57. https://doi.org/10.1016/s2214-109x(18)30408-x.
39. **Wahl B, et al.** (2018) Burden of Streptococcus pneumoniae and Haemophilus influenzae type b disease in children in the era of conjugate vaccines: global, regional, and national estimates for 2000–15. *The Lancet Global Health*; **6**: e744-e757. https://doi.org/10.1016/s2214-109x(18)30247-x.
40. **Chen C, et al.** (2024) Re-evaluating the impact and cost-effectiveness of pneumococcal conjugate vaccine introduction in 112 low-income and middle-income countries in children younger than 5 years: a modelling study. *The Lancet Global Health*; **12**: e1485-e1497.
41. **World Health Organization** (2023) Vaccines could avert half a million deaths associated with anti-microbial resistance a year. Available at https://www.who.int/news/item/28-07-2023-vaccines-could-avert-half-a-million-deaths-associated-with-anti-microbial-resistance-a-year# (accessed January 16 2024).
42. **Chisti MJ, et al.** (2021) Antibiotic-resistant bacteremia in young children hospitalized with pneumonia in Bangladesh is associated with a high mortality rate. In: *Open forum infectious diseases.* Oxford University Press US.
43. **Nguyen KT, et al.** (2021) Pneumonia: drug-related problems and hospital readmissions. *Pneumonia*.
44. **World Health Organization** (n.d.) Vaccines for Antimicrobial Resistance (AMR). Available at https://www.who.int/teams/immunization-vaccines-and-biologicals/product-and-delivery-research/anti-microbial-resistance (accessed October 17 2024).