

**The Deaths Lost in Vaccine Projections: An Examination of the Impact of Vaccination Scenarios on
Forecasted Mortality Attributable to Pneumococcal Lower Respiratory Infections**

Maxwell Pierce

A thesis

Submitted in partial fulfillment of the

Requirements for the degree of

Master of Public health

University of Washington

2022

Committee:

Jon Mosser

Hmwe Kyu

Program Authorized to Offer Degree:

Department of Global Health

©Copyright 2022

Maxwell Pierce

University of Washington

Abstract

The Deaths Lost in Vaccine Projections: An Examination of the Impact of Vaccination Scenarios on Forecasted Mortality Attributable to Pneumococcal Lower Respiratory Infections

Maxwell Pierce

Chair of Supervisory Committee:

Jon Mosser

Department of Global Health

Mortality forecasts can be a way to quantify the impact that causes of death may have over a period of time based on current trends. Cause-specific forecasts depend on associated mitigating factors and drivers. The focus for this work was pneumococcal lower respiratory infections (LRI) and the pneumococcal conjugate vaccine (PCV). The Future Health Scenarios team (FHS) at the Institute for Health Metrics and Evaluation (IHME) uses third-dose PCV (PCV3) forecasts as a covariate for mortality attributable to pneumococcal LRI. Variations in PCV3 forecasting methodologies correspondingly impact the projected mortality attributable to pneumococcal LRI. The goal of this study is to determine the degree of variation in mortality between scenarios attributable to non-vaccination status by estimating mortality arising as a consequence of that status.

Four vaccination coverage forecasts were made, two to serve as the outer bounds of vaccination coverage, and two to represent the scenarios being compared. These forecasts were then used as the basis of forecasts of mortality attributable to pneumococcal LRI and, in turn, the variation in mortality between these vaccination scenarios. Once these data were available, comparisons of total difference in mortality, mortality attributable to vaccination, and ratios of those differences and attributable fractions were made to determine the impact of the vaccination forecasts.

Findings from the work involved in this paper suggest that differences of 1.6% to 2.4% exist between scenarios produced for this work. This translated to between 85 and 114 thousand deaths globally for the under-5 age group attributable to pneumococcal LRI between the years 2022 and 2050. This variation arose solely due to changes in vaccine forecasting methodology with all other variables and covariates held constant between mortality calculations.

Introduction

Lower respiratory infections accounted for approximately 2.5 million (1.18-1.42 95% UI) deaths in 2019, according to estimates produced by the *Global Burden of Disease* study (Kyu et al). The PCV vaccine serves as a protective factor against pneumococcal LRI, serving to reduce associated mortality. Accordingly, PCV3 coverage is a key covariate in the projection of mortality attributable to pneumococcal LRI, as somewhere between 4.7 million and 7.0 million deaths in the under-5 age group between 2022 and 2050 are attributable to pneumococcal LRI, depending on vaccine forecasting methodology used in this paper. As of 2021, PCV3 vaccination coverage globally is approximately 52.0% (Causey et al), which means that there is a great deal of room for growth in vaccination rate, and accordingly room for variation between forecasting methodologies.

Pneumococcal LRI mortality forecasts produced by FHS at IHME use a model that incorporates PCV3 vaccination rate (Foreman et al). By holding all other variables constant, this paper sought to determine how variation in coverage would impact mortality attributable to pneumococcal LRI forecasts to the year 2050. By assessing how mortality varied according to the methodology by which the PCV3 vaccine was forecasted, the consequences of these methods can be better understood in their relationship to mortality.

Methods

This paper based a scenario for vaccine forecasting on methods used by Global Alliance for Vaccines and Immunization (GAVI), the Vaccine Almanac, IHME, and the Vaccine Impact Modeling Consortium (VIMC), organizations which were producing forecasts for the same vaccines at the global level. The VIMC and GAVI both used a similar forecasting methodology, in which coverage increased by a set percentage annually to a maximum of 90% (Li et al). The models used by FHS at IHME base projections of multiple factors, including income, education, fertility, and the coverage of other vaccines (Foreman et al). These two methods are compared to determine how their differences impacted mortality forecasts.

Four vaccine coverage scenarios for PCV3 were produced for this paper: one scenario that reflected 100% future vaccine coverage, one that reflected 0% coverage, one that was based on previous FHS vaccine forecasting methods, and one that was based on methods described by the Vaccine Impact Modeling Consortium, the Vaccine Almanac, and GAVI. For the 100% scenario, future values for vaccination coverage starting in the year 2022 were set to 1. Similarly, for the 0% scenario, future vaccination coverage rates were set to 0. These bounds were chosen to allow an upper and lower bound for vaccine-attributable mortality, with the highest number of deaths occurring in the 100% scenario and the lowest in the 0% scenario. Forecasts began in 2022, with the last past year of data in 2021.

The scenario produced based on methods used by the Vaccine Impact Modelling Consortium and GAVI followed a relatively simple pattern – locations with less than 70% coverage increase in coverage by 3% per year, while locations with more than 70% but less than 90% increased by 1%, finally capping out at 90% (Li et al). This is a simplification of the methods used by these other organizations, but served to incorporate some of the similarities in terms of vaccine coverage increase over time. The primary deviation between methods was in the handling of projected introduction dates. For this paper,

introduction dates were simulated based on covariates including SDI and the coverage of other associated vaccines. For methods used by the VIMC and GAVI, methods for predicting vaccine introductions were based on other factors, such as expert opinion or the date the country in question expected to introduce the vaccine.

For the FHS scenario, vaccine coverage forecasts were produced in accordance with methods used in the paper *“Forecasting life expectancy, years of life lost, and all-cause and cause-specific mortality for 250 causes of death: reference and alternative scenarios for 2016-2040 for 195 countries and territories”* by Foreman et al. For a condensed explanation, the third-dose diphtheria-tetanus-pertussis (DTP3) vaccine is forecasted in logit space as a function of Socio Demographic Index or SDI.

$$\text{logit}(\text{coverage}) = \alpha + \beta * \text{SDI}$$

A ratio of PCV3 to DTP3 is then produced using historical data and forecasted using Spatiotemporal Gaussian Process Regression (ST-GPR) using Health Access and Quality (HAQ) Index and DTP3 coverage forecasts as covariates. The use of DTP3 as a reference allows the use of the more data-rich vaccine to help produce forecasts for long-term PCV3 vaccination rates, as there is comparatively little data for PCV3 as a more recently introduced vaccine. These ratio forecasts are multiplied by DTP3 forecasts to produce final estimates of PCV3 coverage for all nations and subnational states where the PCV3 vaccine has been introduced. In un-introduced locations, vaccination introduction is simulated using SDI, eligibility for GAVI assistance, and coverage of DTP3. Simulations for this process use a Weibull distribution, a form of distribution characterized in this case by an early peak and long tail. This characterization is due to the assumption that, following vaccine development, a few countries with earlier access to these resources will quickly adopt the vaccine, after which the vast majority of countries will adopt fairly quickly. After this, a long tail is projected as more countries without the same degree of access afforded to them more slowly adopt the vaccine. The model used by FHS is fit on past introduction dates, with future introductions predicted by forecasted covariate values.

As there is no past data for PCV3 coverage in locations without introduced PCV3, ratios must be simulated for locations with simulated introductions. This process uses a mixed effects model of Education, DTP3 coverage, and Lag Distributed Income (LDI) (Foreman et al).

$$\text{Logit}(R_{l,y}) = \beta_0 \text{LDI}_{l,y} + \beta_1 \text{YSI}_{l,y} + \beta_2 \text{DTP3}_{l,y} + \pi_0 \text{Edu}_{l,y} + \pi_1_{l,y}$$

R denotes ratio of vaccine coverage, LDI lag-distributed income, YSI years since introduction, DTP3 is associated coverage, and Edu is education. Subscripts l and y denote specificity by location and year, and β refers to fixed effects, while π refers to the random effects used in this model. Figure 1 presents those locations for which PCV3 vaccine has and has not been introduced.

Figure 1: *Locations of PCV3 Introduction as of 2020*

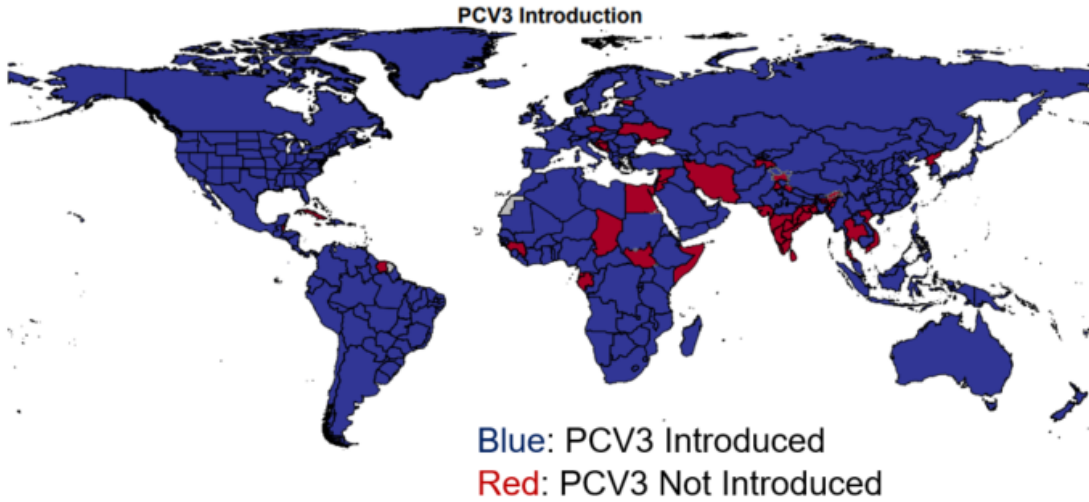


Figure 1: Locations of introduction are blue, whereas unintroduced locations are red.

The same method of mortality estimation was used for every vaccination scenario, and further information can be seen in the methods section of Foreman et al, Lancet 2018. Cause-specific mortality scalars were calculated based on risk factors. In this case, the risk factor is absence of PCV3 vaccination, so a location with 90% vaccine coverage would have 10% exposure to non-vaccinated status. For this method it is assumed that underlying mortality can be attributed to SDI, location, age, and year. A risk factor scalar is calculated based on the ratio of total cause-specific mortality to underlying cause-specific mortality.

$$Scalar = \frac{1}{1 - PAF}$$

Where PAF refers to the population attributable fraction, or the ratio of risk-attributable cause specific deaths. In this case, deaths measured are attributable to pneumococcal LRI due to non-vaccination status. This scalar is used as an offset from the underlying mortality model to account for the impact of vaccination coverage by year.

$$\ln(Total Mortality) = \ln(Underlying Mortality) + \ln (scalar)$$

As underlying mortality is forecasted as a function of SDI, time, age, and location, the final inclusion of the scalar serves to incorporate the effect that results from the associated risk factor.

$$\ln(Total Mortality) = \alpha_{la} + \beta_0 SDI_{<0.8} + \beta_1 SDI_{\geq 0.8} + \gamma_a time + \ln (Scalar)$$

γ_a is a slope fit on age with time as a covariate, while α_{la} is an intercept fit on location and age. Separate values for β are used depending on whether SDI is greater or less than 0.8. All versions, save for vaccine and associated scalar and mortality version, were held constant between runs, so the only variation between mortality forecasts is due to vaccine coverage (Foreman et al).

Once approximations of forecasted death rates attributable to pneumococcal LRI had been produced, they were multiplied by population forecasts produced by the FHS team at IHME. These mortality

forecasts and the variation between them answers the question central to this paper – how many deaths exist in the gaps between vaccine projections.

Results for this paper took the form of age-, sex-, and region-specific mortality attributable to pneumococcal LRI. For the purposes of this investigation and for the sake of brevity, the population of concern was individuals under 5 of all sexes. This meant that this study produced under 5 mortality, or U5M, for all nations, as well as aggregated estimates at the superregion and global levels. In addition, a distinction was made between ‘global’ and ‘all introduced locations’. As introduction date simulation is a source of inconsistency between forecasting methods, analysis on data excluding un-introduced regions was done in order to determine the existence of variation attributable to methodology unrelated to introduction date.

Results

Figure 2: *Global Rate of PCV3 Vaccination*

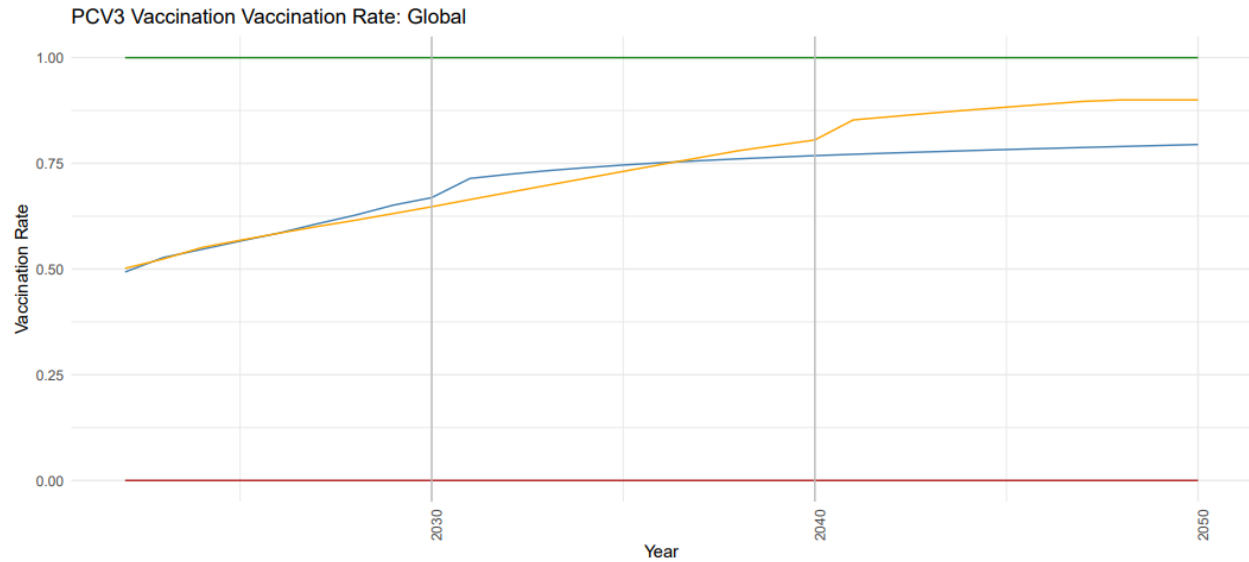


Figure 2: Levels of global vaccination, with 1 representing total vaccination and 0 representing no vaccination. The red scenario represents 0% vaccination coverage, and the green 100%. The orange scenario represents the gradual method of vaccine forecasting, while the blue represents the FHS method. Values may not appear to align with underlying forecasting methodologies, as results here are aggregated from national levels to the global level.

Global aggregates of vaccination forecasts were not used to approximate global mortality and were not produced at the global level. Instead, national vaccination levels were aggregated based on population to reach the global aggregate (Figure 2). While initial forecasts were approximately equal, as both started from the same point in 2021 with a coverage of 52.0%, forecasts diverged to reflect different methodologies, with PCV3 coverage in 2050 at 90% for the Gradual scenario and 79.5% for the FHS scenario.

Figure 3A and 3B

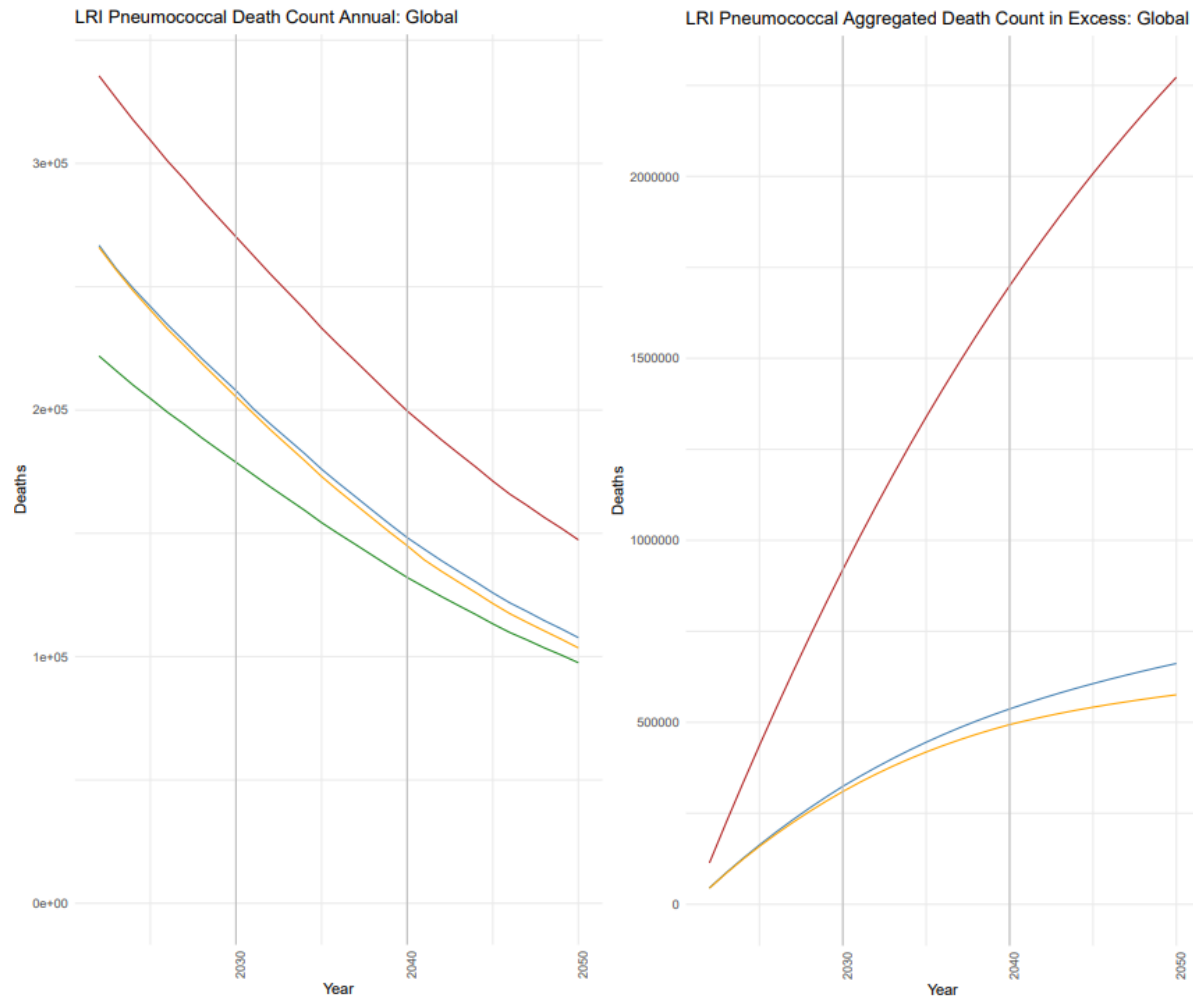


Figure 3: The red scenario represents 0% PCV3 coverage, while the green scenario represents 100% PCV3 coverage. The orange scenario represents the gradual method of vaccine forecasting, while the blue represents the FHS method. On the left are annual counts of death in each location, and on the right are cumulative death counts in excess of the 100% vaccination scenario, which served as a lower limit of mortality in this analysis.

For global and introduced locations respectively, the difference in cumulative death counts between 2022 and 2050 for the FHS and Gradual scenarios was 1.6% and 2.4%. Selecting for locations where the PCV3 vaccine has been introduced, variation between the two scenarios accounted for 2.4% of all pneumococcal LRI deaths in the under-5 age group. This higher rate of variation seems to be due to a tendency for the Gradual scenario to produce higher vaccination and lower associated mortality forecasts in introduced locations when compared to the FHS scenario, and lower vaccination rates and higher mortality rates in un-introduced locations. As can be seen, the variation between the Gradual and FHS scenarios are comparatively small when allowing for the total variation that can exist as a result of vaccination variation. Annual death counts declined globally over time in all scenarios, reflecting improvements in care and general health correlating to other factors improving over time, such as HAQ and SDI, even for the scenario when PCV3 is lowered in forecasts.

Fig 4A: Cumulative Deaths Globally

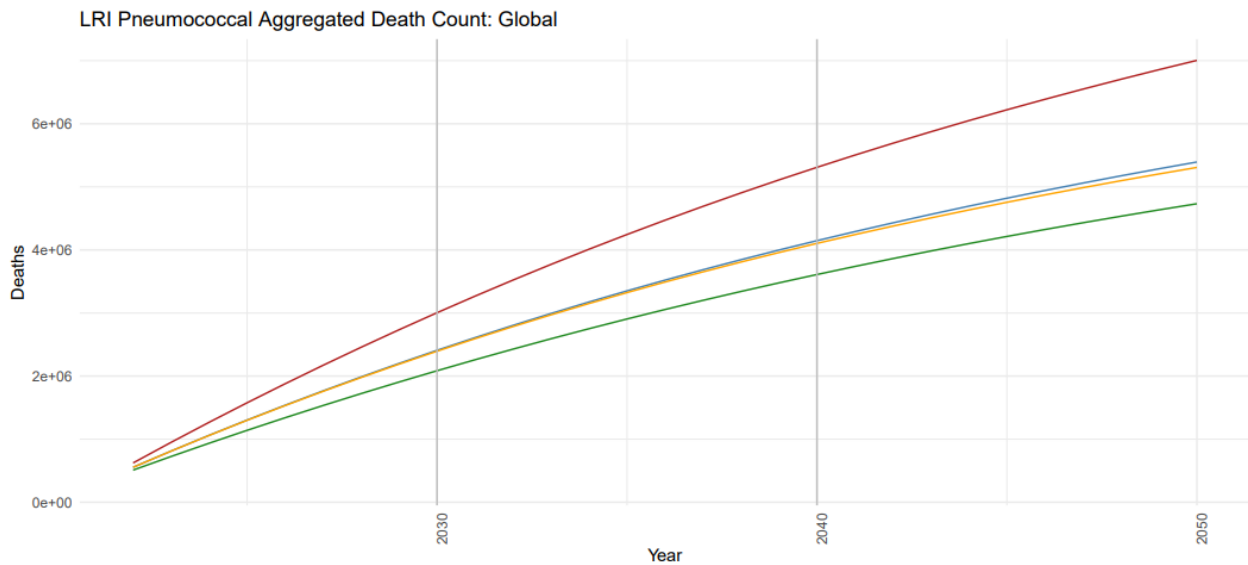


Fig 4B: Cumulative Deaths in Locations with PCV3 Introduced

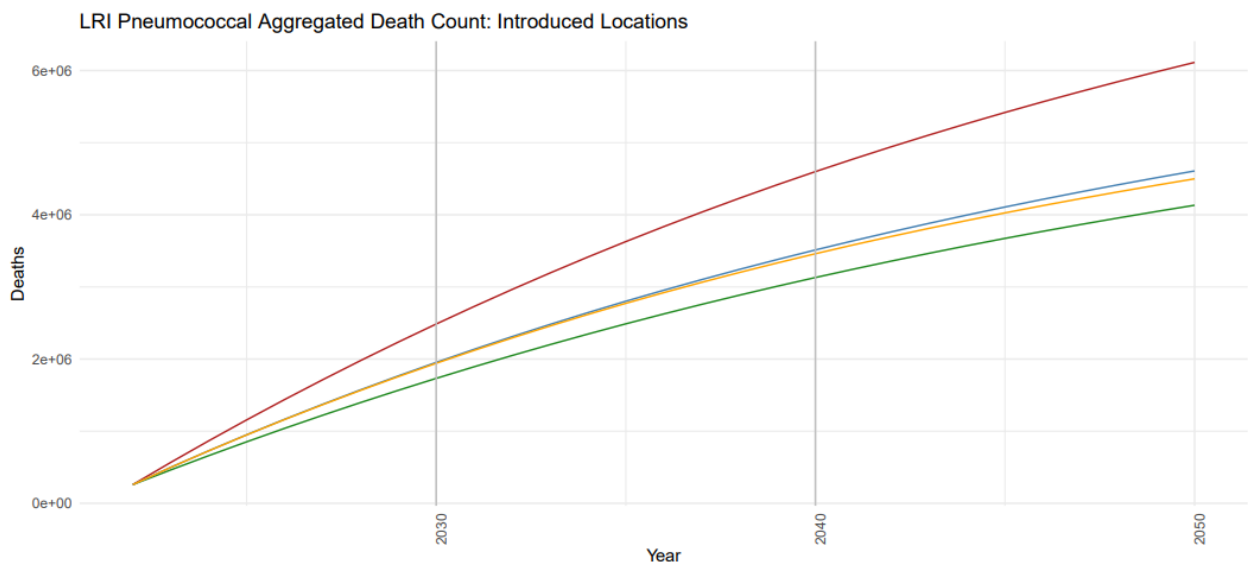


Figure 4: The red scenario represents 0% vaccination coverage, and the green 100%. The orange scenario represents the gradual method of vaccine forecasting, while the blue represents the FHS method. Figure 4A the aggregated cumulative mortality for all locations globally, while figure 4B shows only those locations where PCV3 has been introduced.

There was a total of 7.00 million deaths due to pneumococcal LRI projected globally by the year 2050 under the 0% coverage scenario, compared to 4.73 million under the 100% coverage scenario. The FHS and Gradual scenarios projected 5.39 and 5.31 million deaths respectively, with a difference of

approximately 85 thousand total deaths between them at the global level. However, this variation increased when selecting solely for locations where the PCV3 vaccine has already been introduced, with the discrepancy increasing to approximately 114 thousand deaths by the year 2050 (Table 1).

Table 1: *Cumulative Mortality by Global Regions by 2050 (Thousands)*

	0% Scenario	1% Scenario	FHS Scenario	Gradual Scenario
Global	7,005	4,732	5,394	5,308
Global Excess Mortality	2,272	0	662	576
Introduced Locations Only	6,238	4,216	4,698	4,584
Introduced Locations Excess Mortality	2,023	0	482	368

Table 1: *Deaths (in thousands) attributable to pneumococcal LRI globally and for all introduced locations. Excess mortality refers to mortality greater than the 100% vaccination scenario in each case.*

While the difference may appear relatively small, the distribution of pneumococcal LRI deaths and variation in vaccination is not equitable for either scenario, with higher levels of variation appearing in the Western and Sub-Saharan Africa regions when compared to other global superregions. Those regions with a higher mortality attributable to pneumococcal LRI had a correspondingly higher variation between the FHS and Gradual scenarios.

Table 2: *Cumulative Mortality by Superregion by 2050 (Thousands)*

Superregion	0 Scenario	1 Scenario	FHS Scenario	Gradual Scenario
Southeast Asia, East Asia, and Oceania	407	276	331	325
Central Europe, Eastern Europe, and Central Asia	141	95	100	104
High-income	20	13	14	14
Latin America and Caribbean	191	129	140	139
North Africa and Middle East	326	220	244	250
South Asia	1,087	741	880	857
Sub-Saharan Africa	4,833	3,257	3,684	3,618

Table 2: Deaths (thousands) attributable to pneumococcal LRI in superregions as defined by IHME's Global Burden of Disease study.

Discussion

A variation of between 1.6% and 2.4% of projected global mortality in the under-5 age group attributable to pneumococcal lower respiratory infections exists between the methods discussed in this paper, a gap of approximately 114 thousand deaths of children less than five years old. This arose from a difference of 10% vaccine coverage by the year 2050 between forecasting methodologies examined. The mortality variance attributable to PCV3 coverage demonstrates as even fairly minor changes in vaccine forecasts can account for differences of many thousands of deaths when applied to only a single vaccine and associated cause.

The question of vaccine projection is not simply a global or theoretical issue – it is a question of equity in public health as well. Sub-Saharan Africa alone accounts for over 3 million pneumococcal LRI deaths in the under-5 age group by the year 2050. Between the FHS and Gradual scenarios was a difference of approximately 66 thousand deaths. Vaccine projections are typically produced in countries for whom vaccination against many major causes has been effectively completed. However, for those countries for whom the difference in vaccination method has a significant impact, this difference can be very meaningful in terms of death and should be taken seriously.

Several limitations for this paper arose due to either limited resources, limited information, or decisions to prioritize other issues. Ultimately, producing and analyzing data for both vaccines proved to be too large a task, so one was removed. Similarly, production of better and worse mortality projections associated with each scenario to serve as a form of confidence interval for the FHS models was not performed due to time and resource constraints. These decisions limit the interpretability of the forecasts themselves, as there is likely to be some amount of overlap when confidence intervals are included. An additional modelling related challenge that presented itself was the issue of introduction dates for un-introduced locations, as different organizations varied vastly in approach. The inclusion of data relating to only introduced locations was an attempt to demonstrate results even without accounting for those locations, and did manage to show a higher variation between scenarios of 114 compared to 85 thousand deaths. Even beyond dealing with introduction simulation, there are differences between scenarios significant enough to account for over 2.4% of total attributable mortality.

The goal of this paper was to examine what variation exists between methods of vaccine forecasting in the projected impact of mortality. This paper is not an exhaustive analysis – other modes of mortality projection that use covariates as drivers should be tested. A single mortality model was used to maintain consistency, but further work using other models with varied forecasting methodologies is worth further examination and testing. Results, while small on the scale of worldwide mortality, are not representative of all causes and all vaccines - pneumococcal lower respiratory infections represent a comparatively small fraction of deaths in the under-5 age group when compared to other causes such as rotavirus. Future work with other models of mortality could help to validate these results, as well as the inclusion

of other vaccines and causes to examine the total impact attributable to vaccine forecasting methodology. As covariate forecasting and mortality modelling methods vary between organizations and researchers, an understanding of how covariate differences drive those models should be seriously considered and researched.

Vaccination as a driver of mortality can have a meaningful and lasting impact on mortality in the future. Even comparatively minor variations in vaccination can account for thousands of deaths by a single cause when sufficient time is allowed. The differences between these methods of vaccine projection is deserving of additional research and more effort should be given in the future to properly model these drivers of mortality in order to better understand not only what the future of mortality holds, but [also](#) what we can expect the impact of these drivers to be.

References

- KJ Foreman, N Marquez, A Dolgert, et al., Forecasting life expectancy, years of life lost, and all-cause and cause-specific mortality for 250 causes of death: reference and alternative scenarios for 2016–40 for 195 countries and territories, *Lancet* 2018
- SE Vollset, E Goren, CW Yuan, *et al.* Fertility, mortality, migration, and population scenarios for 195 countries and territories from 2017 to 2100: a forecasting analysis for the Global Burden of Disease Study, *Lancet* 2020
- Li X, Mukandavire C, Cucunubá ZM, et al. Estimating the health impact of vaccination against ten pathogens in 98 low-income and middle-income countries from 2000 to 2030: a modelling study [published correction appears in *Lancet*. 2021 Feb 20;397(10275):670]. *Lancet*. 2021;397(10272):398-408. doi:10.1016/S0140-6736(20)32657-X
- Toor J, Echeverria-Londono S, Li X, et al. Lives saved with vaccination for 10 pathogens across 112 countries in a pre-COVID-19 world. *Elife*. 2021;10:e67635. Published 2021 Jul 13. doi:10.7554/eLife.67635
- HH Kyu, SI Hay, M Brauer, et al. “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.” *The Lancet. Infectious diseases*, S1473-3099(22)00510-2. 11 Aug. 2022, doi:10.1016/S1473-3099(22)00510-2
- Causey K, Fullman N, Sorensen RJD, et al. Estimating global and regional disruptions to routine childhood vaccine coverage during the COVID-19 pandemic in 2020: a modelling study. *Lancet*. 2021;398(10299):522-534. doi:10.1016/S0140-6736(21)01337-4