UNIVERSITAT DE VALÈNCIA
FACULTAT DE MATEMÀTIQUES

DEPARTAMENT D´ESTADISTICA I INVESTIGACIÓ OPERATIVA

Bayesian models for the analysis of infectious diseases and immunization strategies

Mónica López Lacort

DOCTORAL THESIS UV/ 03-2024

THESIS SUPERVISORS:
Dra. Ana Corberán Vallet
Dr. Francisco José Santonja Gómez

" I simply wish that, in a matter which so closely concerns the wellbeing of the human race, no decision shall be made without all the knowledge which a little analysis and calculation can provide."

Daniel Bernoulli (1700–1782)

To my family and friends.

Acknowledgements (Spanish)

Gracias a mis directores de tesis, mis “tíos” Ana y Francisco, por su orientación y dedicación a este proyecto. Agradezco sinceramente su paciencia, su guía, sus valiosas sugerencias y sus ánimos en los momentos más difíciles. También quiero extender mis agradecimientos a Álvaro, su implicación ha sido de gran ayuda para poder finalizar la tesis. Que sigamos “tomando nota” por muchos años más.

A mis hermanos científicos, Cintia y Alex, gracias por su inestimable contribución a esta tesis. Pero más allá de eso, las risas que compartimos, la amistad y la confianza que tenemos hacen que trabajar con vosotros sea una experiencia excepcional. Cintia, eres un pilar fundamental en mi vida, gracias, por tanto. Alex, gracias por estar siempre cerca, especialmente cuando volamos juntos.

A Javier, gracias por tu confianza, sabiduría y por todos los conocimientos impartidos. Gracias por dejarme la libertad de aplicar mi creatividad y por propiciar un entorno de trabajo enfocado a la innovación. Gracias por enseñarme a hacer ciencia.

El agradecimiento más merecido es para mi marido, esta tesis es tan tuya como mía, gracias por acompañarme y ser mi impulso en todas mis inquietudes, por ayudarme a crecer, por quererme tan bien y por regalarme ese tiempo tan valioso para que haya podido cumplir esta meta. Compartir la vida contigo, Miguel, es una pasada. Te admiro y te quiero.

A mi hija Carmen, en mitad de la tesis, llegaste para presentarme a una nueva Mónica. Gracias por enseñarme a amar de la manera más pura, por darle mayor intensidad a mis emociones y por mostrarme lo que realmente importa en la vida. Aunque en algunos momentos sentí que te estaba robando tiempo, me di cuenta de que solo era un préstamo; espero que los días en los que mamá estaba ocupada te sirvan como ejemplo, para que seas una luchadora y persigas tus sueños. Recuerda, las cosas que requieren esfuerzo son las que realmente valen la pena. Te quiere mami.

A mi padre, gracias por ser mi ejemplo en el trabajo duro, en el espíritu de superación y en ser buena persona. Gracias por estar siempre, por tus sabios consejos y por tu amor

incondicional. Eres la persona más bondadosa que conozco y estoy muy orgullosa de tenerte como padre.

A mi madre, aunque no estés físicamente conmigo siempre te siento cerca. Eres mi guía y tu amor está presente en este logro. A mi cielo, os echo muchísimo de menos, pero sé que me habéis estado cuidando y dando fuerzas.

Gracias a mi hermano, a Isabel, y a todos mis familiares y amigos, por ser un gran apoyo. Vuestra presencia y ayuda han sido fundamentales en mi camino. Isabel, muchas gracias por cuidar tan bien de Carmen y por estar disponible en cualquier momento que te necesité para poder dedicarme a la tesis.

A mi surmana Sandra, por todas esas citas a las 9 que me han llenado de energía, gracias por tu apoyo incondicional en cada paso de este viaje. A Patri, Irene y Bea, por estar ahí siempre. A Noelia y Mónica, gracias por nuestros encuentros vitamina, sois especiales. A mi Ainara, trabajar contigo es maravilloso, pero tu amistad es insuperable. A Bego, gracias por tantos momentos compartidos, fuiste la primera en acogerme en el AIV, y eso significa mucho para mí.

A todos mis compañeros del Área de Investigación en Vacunas que me habéis acompañado en esta etapa: gracias por hacer de cada día en el AIV una oportunidad para aprender y disfrutar. Sois mi familia científica.

Abstract:

This thesis aims to advance Bayesian modelling approaches for comprehending infectious disease patterns and epidemics, with a specific focus on contributing methodological enhancements to immunization and vaccination research. The research presented first explores Bayesian temporal models in discrete time to describe disease incidence. Then, it delves into spatiotemporal Bayesian models for cluster detection and identification of small areas that follow a similar trend. Discrete-time models are applied to two case studies for the study of bronchiolitis and the respiratory syncytial virus (RSV), while the spatiotemporal approach is applied to a different case study corresponding to human papillomavirus (HPV) vaccination. These case studies address important questions in the field of vaccine research. Namely, in the first case study, a Bayesian stochastic model is developed to describe bronchiolitis dynamics in young children. The proposed model is a multivariate age-structured model that accurately captures both endemic and epidemic periods. It considers interactions among age groups by combining ideas from compartmental models and Bayesian hierarchical Poisson models or negative binomial models in a novel way. The study employs high-quality population-based data from the Valencian Integrated Databases (VID). In a second stage, the proposed model is extended to simulate the potential effects of different newborn immunization scenarios against the RSV on bronchiolitis. We provide an app tool that estimates the expected reduction in bronchiolitis episodes for a range of different values of the uptake and effectiveness.

The second case study focuses on the recent approval of the monoclonal antibody nirsevimab and the bivalent prefusion RSV vaccine for maternal immunization. We employ an adaptation of the previously proposed model-based approach to estimate a realistic reduction of RSV bronchiolitis in infants under 12 months old if each of these two interventions is applied. This study aims to have a greater knowledge of the disease and more precise and realistic impact estimations considering new findings regarding efficacy and protection duration in clinical trials. Results indicate that both interventions could effectively reduce RSV bronchiolitis and hospitalizations, providing valuable insights into their potential impact.

The third case study aims to detect suboptimal vaccination coverage of the HPV vaccine in the Valencia Region, despite its inclusion in the systematic vaccination program since

2008. A spatiotemporal clustering model is developed to identify health districts with similar vaccination behaviour. The model allows for the exploration of different spatial and temporal structures, revealing suboptimal HPV vaccination coverage in specific health districts. This information is essential for implementing targeted strategies to enhance vaccination coverage.

In conclusion, this thesis contributes methodological improvements to the field of infectious disease modelling, disease mapping and vaccination research. The developed Bayesian models offer flexible and effective tools for understanding disease dynamics and assessing the potential impact of immunization strategies. The case studies provide valuable insights for public health decision-making and highlight the importance of considering real-world data in modelling studies.

Resumen:

Esta tesis tiene como objetivo avanzar en la modelización bayesiana para comprender la dinámica de enfermedades infecciosas, con especial hincapié en proporcionar metodología adecuada para la investigación del efecto de estrategias de inmunización y vacunación. La investigación presentada explora, en primer lugar, modelos temporales bayesianos en tiempo discreto para describir la incidencia de enfermedades infecciosas. A continuación, profundiza en el desarrollo de modelos bayesianos espaciotemporales para la detección de conglomerados e identificación de áreas pequeñas que siguen una tendencia similar. Los modelos temporales son aplicados a dos estudios de caso: la bronquiolitis y el virus sincitial respiratorio (VSR), mientras que el modelo espaciotemporal se aplica a un estudio de caso diferente correspondiente a la vacunación contra el virus del papiloma humano (VPH). Estos estudios de caso abordan preguntas fundamentales en el campo de la investigación de vacunas. Específicamente, en el primer estudio de caso, se desarrolla un modelo estocástico bayesiano para describir la dinámica de la bronquiolitis en niños pequeños. El modelo propuesto es un modelo multivariado estructurado por edades que captura con precisión tanto los períodos endémicos como epidémicos. Además, considera las interacciones entre grupos de edad combinando ideas de modelos compartimentales y modelos jerárquicos bayesianos de Poisson o modelos binomiales negativos de una manera novedosa. El estudio emplea datos poblacionales de alta calidad de las Bases de Datos Integradas de la Comunidad Valenciana. En una

segunda etapa, el modelo propuesto se extiende para simular los efectos potenciales de diferentes escenarios de inmunización neonatal contra el VSR en la bronquiolitis. Para este estudio proporcionamos una aplicación que estima la reducción esperada en episodios de bronquiolitis para diferentes combinaciones de la cobertura y efectividad.

El segundo estudio de caso se centra en la aprobación reciente del anticuerpo monoclonal nirsevimab y la vacuna bivalente de perfusión contra el VSR para la inmunización materna. Empleamos una adaptación del modelo desarrollado previamente para estimar de manera realista la reducción de la bronquiolitis por VSR en lactantes menores de 12 meses si se aplican cada una de estas dos intervenciones. Este estudio tiene como objetivo obtener un mayor conocimiento de la enfermedad y estimaciones de impacto más precisas y realistas considerando los nuevos resultados sobre eficacia y duración de la protección en ensayos clínicos. Los resultados indican que ambas intervenciones podrían reducir efectivamente la bronquiolitis por VSR y las hospitalizaciones, proporcionando información valiosa sobre su impacto potencial.

El tercer estudio se centra en la detección de coberturas de vacunación subóptimas contra el VPH en la Comunidad Valenciana, a pesar de su inclusión en el programa de vacunación sistemática desde 2008. Para ello, desarrollamos un modelo de agrupamiento espaciotemporal, que nos permite identificar distritos sanitarios con comportamientos de vacunación similares. El modelo permite explorar diferentes estructuras espaciales y temporales, revelando una cobertura de vacunación contra el VPH subóptima en distritos sanitarios específicos. Esta información es esencial para implementar estrategias dirigidas para mejorar la cobertura de vacunación.

En conclusión, esta tesis aporta mejoras metodológicas en el campo de la modelización de enfermedades infecciosas, mapeo de enfermedades e investigación en vacunas. Los modelos bayesianos desarrollados ofrecen herramientas flexibles y efectivas para comprender la dinámica de estas enfermedades y evaluar el impacto potencial de las estrategias de inmunización consideradas. Los estudios de caso proporcionan conocimientos valiosos para la toma de decisiones en salud pública y destacan la importancia de considerar datos del mundo real en los estudios de modelización.

Thesis structure:

This thesis is structured into five chapters. Chapter one contextualizes the theme of the thesis. It explains key concepts such as infectious diseases, epidemics, vaccination, and the role of mathematics and statistics in epidemic prevention and vaccination research. We then delve into key concepts regarding Bayesian statistics and commonly used models for epidemic modelling and vaccination research, focusing on those that have motivated our research.

In the second chapter, we present the first case study, delving into discrete-time stochastic models. Specifically, we introduce a Bayesian multivariate stochastic age-structured model for the study of bronchiolitis dynamics. An extension of the methodology is also presented to simulate the potential impact if a proportion of newborn were immunized.

In the third chapter, we apply a modified version of the previously described model to gain more insight into RSV dynamics. In particular, the model includes finer age groups and considers the latest findings from clinical trials. This model better simulates the impact of two immunization strategies on RSV.

In the fourth chapter, we present a novel spatiotemporal clustering model for the study of the HPV vaccination coverage. The model allows exploring different spatial and temporal structures, revealing suboptimal HPV vaccination coverage in specific health districts.

These three chapters follow the same structure. They include an abstract followed by an introduction, which motivates the study and incorporates a review of related literature. Details about the proposed methodology and main results are then provided. At the end, some concluding remarks and directions for future research are given.

The last chapter offers general concluding remarks as well as key insights and avenues future research.

Table of contents:

List of tables:

List of figures:

Chapter 1

Introduction

1. With Bayes, Bernoulli, and Jenner, it all begins.

Nowadays, we can use Bayesian models to understand the dynamics of infectious diseases and to either describe or simulate the effects of vaccination programs and other forms of immunization. This is possible thanks to three pivotal discoveries at the end of the eighteenth century: the revelation of Bayes' theorem, the formulation of the first epidemiological model, and the development of the first vaccine. These historical milestones laid the foundation for our current understanding of the interconnectedness between mathematics, epidemiology, and vaccines.

The history of mathematics behind epidemics begins with two significant breakthroughs during the smallpox era, a period marked by the prevalence of one of the most devastating diseases in the humanity caused by the variola virus. In 1766, Daniel Bernoulli published a groundbreaking paper titled "Essai d'une nouvelle analyse de la mortalité causée par la petite vérole", presenting the inaugural mathematical model for analysing the spread of smallpox. Bernoulli argued the benefits of variolation, an inoculation technique exposing individuals without prior smallpox experience to material from smallpox sores (pustules). In 1772, D'Alembert further advanced Bernoulli's work by incorporating age-dependent parameters into the model. While Laplace also made contributions to the concept, systematic development only materialized with the 1911 benchmark paper by Ross, establishing the foundations of modern mathematical epidemiology (1–4).

Twenty years after, in 1796, Edward Jenner discovered the smallpox vaccine, marking a vital moment in medical history. His merit is underscored by the practical evidence he presented, demonstrating that the inoculation of material from a person with cowpox lesions to a healthy child can confer protection against smallpox. This groundbreaking discovery, which laid the foundation for modern vaccination, transpired in Europe between the late eighteenth and early nineteenth centuries. Jenner's pioneering work not

only transformed the approach to disease prevention but it also set the stage for subsequent advancements in immunization science (5).

The other key discovery emerged simultaneously to Bernoulli's model in 1763. The Bayes' theorem, formulated by Thomas Bayes, was presented two years after his death by the philosopher Richard Price to the Royal Society. Thomas Bayes' manuscript, titled 'An Essay Towards Solving a Problem in the Doctrine of Chances', provided the first clear solution to the problem of inverse probability. It describes how we can calculate the probability of the occurrence of an event given the known probability of a certain condition. The theorem was abandoned until 1812, when it was independently rediscovered by Pierre Simon Laplace. There is a belief that Laplace might not have been aware of Bayes' theorem, as he developed a more formal version of it during his rediscovery. Until the end of the 20th century, the Bayesian approach was in disuse due to time requirements. New technological advances have recently increased the use of Bayesian models in many fields, including epidemiology and for public health decision-making (6). Nowadays, epidemics models, vaccination research and Bayesian analysis are intimately connected.

2. Infectious diseases, immunization, mathematics and beyond

This section introduces key concepts that are essential for contextualizing and understanding the thesis: infectious diseases, vaccination as a preventive measure, and the role of mathematical models in public health decision-making.

Infectious diseases are disorders caused by various microorganisms like viruses, bacteria, protozoa, and fungi. They can be transmitted from person to person, from animals to humans, or from the environment to humans. They represent the leading cause of morbidity and mortality worldwide, having more impact in middle- and low-income countries (7). Infectious diseases exhibit various dissemination patterns: epidemic, pandemic, endemic, and outbreak. An epidemic denotes the sudden and widespread occurrence of a disease, exceeding expected cases, in a specific population or region. When an epidemic spans international borders and affects a significant global population, it becomes a pandemic. In contrast, endemic signifies the persistent, usual presence of a disease within a specific region or population, indicating a steady baseline prevalence.

An outbreak, like an epidemic but more localized, is related with an increased number of cases in a specific community (8). In today's interconnected world, the COVID-19 pandemic serves as prime evidence of the global impact that epidemics and pandemics can have. Health crises not only affect specific regions, but also have far-reaching social and economic consequences that transcend geographical boundaries (7,9,10).

Several infectious diseases, such as COVID-19, measles, or rotavirus infection, can be prevented by vaccination, a simple, safe, and cost-effective way of protection. Vaccines activate artificially the immune system to build resistance to specific infections and make the immune system stronger (11). In addition to the direct protection offered to vaccinated individuals, vaccines also have indirect effects by curbing disease transmission from the vaccinated to others (12,13). Ensuring adequate levels of herd immunity is the only reliable method for preventing epidemics and a re-emergence of vaccine-preventable diseases (VPDs) (14). Herd immunity is achieved when a significant portion of a population becomes immune to an infectious disease, either through vaccination or prior illness, reducing so the likelihood of person-to-person transmission. This protection extends to individuals who are not vaccinated, such as newborns and those with chronic illnesses, as the disease finds limited opportunities to spread within the community (15).

Other ways of protection, such as passive immunity, are also available. It occurs when a person receives antibodies to a disease rather than producing them internally. This immediate protection can be acquired from sources such as the mother through the placenta or through antibody-containing blood products. Unlike active immunity, passive immunity is short-lived, lasting only for a few weeks or months (16). Another avenue for passive protection is the use of monoclonal antibodies (mAbs), which are laboratory-produced molecules designed to mimic the immune system's ability to fight off harmful pathogens, such as viruses or bacteria They are particularly beneficial for immunosuppressed individuals. They also present advantages in generic manufacturing, facilitating a swift response during outbreaks (17).

Public health agencies play a crucial role in the detection, prevention, and control of infections within the population. They also have the responsibility to implement and assess immunization programs to ensure sufficient protection. During the second half of the 20th century, vaccination programs have successfully eradicated smallpox, nearly

eliminated polio, and substantially reduced viral and bacterial infections in children through national immunization programs in developed countries. It is considered the most effective medical intervention, in fact, vaccines have prevented nearly one-quarter (21.7%) of the 5.3 million deaths among children under the age of 5 years in 2019 (18). However, outbreaks of VPDs continue to occur even in countries with well-established vaccination programs. Reasons include the existence of under-vaccinated populations due to the increasing anti-vaccination movement, the increasing movement of individuals across borders, or socioeconomic inequalities in the population (14,19).

Mathematical and statistical models play a pivotal role in epidemiology, serving as indispensable tools for understanding the dynamics and key features of infectious diseases. Beyond their scientific significance, these models contribute to shaping political decisions and guiding interventions. For instance, they allow analysing early outbreak stages, assessing interventions like immunization programs, identifying risk factors related to either individuals or the environment and analysing spatial and temporal patterns. Moreover, mathematical and statistical models facilitate hypothesis generation and formal testing, effectively minimizing bias and enhancing the precision of epidemiological observations (6,20–22). The influence of these models is growing due to the increasing abundance of data on infectious diseases, a consequence of the advancements in technology and the rise of big data in health data collection and infectious disease surveillance (23,24).

As illustrative examples, during the 2009 influenza pandemic, mathematical models assisted the World Health Organization (WHO) and national response units in interpreting outbreak data and guiding decisions on vaccination strategies. They were used to estimate the basic reproduction number, represent the average number of secondary infections per infected individual in a susceptible population, and evaluate the impact of vaccination timing and targeting on the epidemic's peak and duration. Another example is the 2014-2016 Ebola outbreak in West Africa. In this case, mathematical models helped estimating key parameters for outbreak control measures, including case isolation and contact tracing. The experience with Ebola contributed to the development of guidance for designing vaccine efficacy trials during public health emergencies (25). The most recent example is the multitude of mathematical models developed during the COVID-19 pandemic, providing insights into contagion dynamics and response strategies

(26). Bayesian applications in epidemiology, such as the spatiotemporal cluster detection of varicella (27), the assessment of the impact of rotavirus vaccination on hospitalizations due to rotavirus (28), and the evaluation of the impact of temperature on the transmission of Zika, dengue, and chikungunya (29), showcase the versatility of Bayesian methods for understanding and addressing complex issues in public health.

Models can be classified based on their purpose: explanatory, descriptive, or predictive. Explanatory modelling is commonly used for causal testing. Descriptive modelling is the most utilized and developed by statisticians. Descriptive models aim to summarize or represent data structures in a concise manner. Unlike explanatory modelling, these models do not heavily rely on an underlying causal theory. They focus on the measurable level rather than the construct level. A regression model is descriptive when it captures associations between dependent and independent variables without emphasizing causal inference or prediction. Predictive modelling involves applying statistical models or data mining algorithms to predict new or future observations (30).

3. Bayesian statistics, key concepts.

Two approaches for statistical analyses are the frequentist paradigm and the Bayesian paradigm. The influence and utilization of Bayesian statistics have expanded significantly across diverse facets of medical research and practice, mainly in epidemiology research (31). A key distinction lies in their treatment of parameters and the interpretation of the probability. The frequentist approach considers the parameters as fixed quantities to be determined. In the Bayesian approach, the model parameters are assumed to be random variables, and so we have to determine their distribution. Unlike the frequentist approach, Bayesian inference allows the formal inclusion of prior knowledge in the analysis, combining information from the observed data with the researcher's accumulated experience through specified prior probabilities. Furthermore, Bayesian methods excel in statistical inference for complex models that pose challenges for frequentist techniques. While frequentist approaches may encounter difficulties in estimating parameters in such cases, Bayesian methods provide a robust solution by simulating samples from the posterior distribution of the parameters of interest using Markov chain Monte Carlo (MCMC) techniques (6).

3.1 Parameter estimation

In the Bayesian paradigm, parameters are treated as random variables, and so they are given a prior distribution. Once the information from the data is incorporated, this prior distribution is updated using Bayes' theorem, obtaining so the posterior distribution:

$$P(\theta|Y) = \frac{P(\theta)P(Y|\theta)}{P(Y)} = \frac{P(\theta)P(Y|\theta)}{\int_{\Omega} P(\theta)P(Y|\theta)d\theta} \propto P(\theta)P(Y|\theta),$$

where:

θ is the set of model parameters.

P(θ) is the prior distribution; that is, before observing the data. If prior information is unavailable, the priors considered are non-informative, and all the information will be obtained from the data.

P(Y|θ) is the likelihood function quantifying the information provided by the data about θ.

P(θ|Y) is the posterior distribution; that is, the distribution of θ once the information from the data is incorporated.

In Bayesian inference, the primary focus revolves around properties derived from the posterior distribution of the model parameters or related functions. These properties are often expressed as expectations computed over the posterior distribution P(θ|Y):

$$E(g(\theta)|Y) = \int_{\Omega} g(\theta)P(\theta|Y)d\theta$$

On many occasions, the posterior distribution is analytically intractable. A possible solution to estimate the quantities of interest is the use of MCMC simulation methods, which involve simulating samples from the posterior distributions (32). Two commonly used MCMC simulation techniques are the Gibbs sampling procedure and the Metropolis-Hastings algorithm.

3.2 Gibbs sampling and Metropolis-Hastings

One of the most popular MCMC methods is the Gibbs sampling procedure, which involves sampling from the known conditional distributions. Let Y be the data vector, θ = (θ_1 , θ_2, , ..., θ_k) the set of parameters, and (θ_1^0, θ_2^0, , ..., θ_k^0) an initial value for the parameters. The procedure proceeds as follows:

$$\theta_1^{(t+1)} \sim P(\theta_1|Y, \theta_2 = \theta_2^{(t)}, \dots, \theta_k = \theta_k^{(t)})$$

$$\theta_2^{(t+1)} \sim P(\theta_2|Y, \theta_1 = \theta_1^{(t+1)}, \theta_3 = \theta_3^{(t)} \ \dots, \theta_k = \theta_k^{(t)})$$

$$\vdots$$

$$\theta_k^{(t+1)} \sim P(\theta_k|Y, \theta_1 = \theta_1^{(t)}, \dots, \theta_{k-1} = \theta_{k-1}^{(t)})$$

for t = 0, ..., N, N being the number of simulated values from the posterior distribution (33).

Another commonly used method for MCMC is the Metropolis-Hastings (MH) algorithm. Here, a proposal distribution is used to generate candidate values for the parameters. In particular, the method generates $\theta' = (\theta_1', \theta_2', \dots, \theta_k')$ values from the proposal distribution $Q\left(\theta \mid \theta^{(t)}\right)$ given the parameter $\theta^{(t)}$. The acceptance of this new vector of parameter values depends on an acceptance rate defined as:

$$\alpha\left(\theta'\middle|\theta^{(t)}\right) = min\left(1, \frac{P(\theta'|Y)Q\left(\theta^{(t)} \mid \theta'\right)}{P\left(\theta^{(t)}\middle|Y\right)Q\left(\theta' \mid \theta^{(t)}\right).}\right)$$

For the acceptance-rejection decision, a number u is generated from a U (0,1) distribution. If $u \leq \alpha$ the proposed candidate is accepted; that is, $\theta^{(t+1)} = \theta'$; otherwise, $\theta^{(t+1)} = \theta^{(t)}$.

Gibbs sampling is a special case of MH sampling where proposal distributions are the posterior conditionals. Recall that all proposals are accepted in Gibbs sampling, which implies that the acceptance probability is always 1 (34).

3.3 MCMC diagnoses

When running an MCMC algorithm, it is important to examine whether the simulated values have approximately converged to the stationary distribution P(θ|X). When implementing the model, it is common to launch two or three chains with N simulations for each parameter. To assess convergence, the simulated values $\theta^{(t)}$ are plotted against t (t = p, …, N) for each of the simulated chains. Here, p is the number of simulations in the burn-in period. These initial simulations are not considered for posterior estimation as they belong to the chain adaptation process and are typically highly correlated. If the

chains cross each other and exhibit a stationary behaviour, it is indicative that they have good convergence and, consequently, the conducted simulation provides a reliable estimate of the posterior (35). Other measures such as the effective sample size (neff) or Rhat are commonly used for MCMC diagnostics, with neff values above 100 and an Rhat around one indicating good convergence.

3.4 Bayesian software

The field of Bayesian inference has seen significant advancements in software development, with a range of tools and libraries catering to different needs and preferences. In this thesis, we used WinBUGS (Bayesian inference Using Gibbs Sampling) and Nimble, which is an extension of the BUGS language but also implements a modelling language embedded in R. Several Bayesian computation methods are implemented, including Metropolis Hastings, Gibbs sampling, and sequential Monte Carlo among others (35).

4. Commonly used models for infectious diseases epidemics and vaccination research

A high variety of mathematical and statistical models as well as computational tools have been proposed both to analyse the transmission and spread of diseases and the effects of interventions and to anticipate the future course of epidemics under different scenarios (36).

4.1 Mathematical modelling

4.1.1 The general epidemic model

Compartmental models are widely used to study the transmission dynamics of infectious diseases or to calculate the percentage of people that have to be vaccinated for herd immunity. One of the simplest compartmental models is the 'S-I-R' model, also known as the general epidemic model, where the population is divided into three compartments based on their epidemiological status. Here, S(t) represents susceptible individuals, I(t) represents the infected, and R(t) represents individuals removed by recovery or death at time t. The total population at time t, N(t), is given by the sum of the susceptible, infected, and removed individuals ($N(t) = S(t) + I(t) + R(t)$). For simplicity in notation, we will refer to S(t) as S, I(t) as I, and so forth. The rates of transfer between compartments are

expressed mathematically as derivatives with respect to time of the sizes of the compartments. The solution to the model can be obtained through (partial) differential equations:

$$\frac{dS}{dt} = -\beta.\frac{S.I}{N}$$

$$\frac{dI}{dt} = \beta.\frac{S.I}{N} - \gamma.I$$

$$\frac{dR}{dt} = \gamma.I$$

where the initial values of the epidemic are S(0) = N-1, I(0) = 1, and R(0) = 0 . Parameter β is the effective contact rate (reflects both the average of contacts per time period and the probability of disease transmission per contact) and γ is the recovery rate. The increase in disease incidence is observed if $\frac{dI}{dt} > 0$; that is, if $\beta.\frac{S.I}{N} > \gamma.I$ or, equivalently, if $\frac{\beta S}{\gamma N} > 1$. Note that if all the individuals are susceptible to be infected, then N is equal to S. In that case, an increase in disease incidence will be observed if $\frac{\beta}{\gamma} > 1$. This quantity is known as the basic reproduction number, $R_0 = \frac{\beta}{\gamma}$, and describes the number of secondary infections produced by one case during his infectious period in a completely susceptible population. If R_0>1, the infection can spread; once a significant portion of the population becomes immune, the likelihood of infected individuals encountering susceptible ones decreases, reducing the average number of secondary cases. If R_0<1, the epidemic is under control (1,36,37).

Vaccination plays a crucial role in controlling disease outbreaks. If a fraction v of the population is vaccinated before an outbreak, the number of infections caused by an individual decreases to $R_0 \cdot (1-v)$. This is because only 1−v of all contacts result in infection. The new reproduction number is then given $R_v = (1-v) \cdot R_0$. If R_v>1, a positive fraction of the community may get infected. This condition is equivalent to $v<1-1/R_0$. The critical vaccination coverage, or the fraction necessary to vaccinate, is denoted $v_c = 1-1/R_0$ (38).

This general epidemic model can be adapted according to relevant characteristics that can influence disease transmission, such as the age, vaccination status, sex, or region, among others. This can be done by partitioning the compartments based on these

characteristics. Also, different tools have been proposed to compute R_0 as a function of other parameters (36).

Depending on the nature of the disease under study, other compartmental models have been formulated: the SIS model describes a disease with no immunity against re-infection; the SEIR or SEIS models include an exposed period between becoming infected and being infective; the SIRS model captures temporary immunity on recovery from infection; the SVIR model includes a compartment accounting for vaccinated individuals (39,40).

For more complex models with many partitions of the population, this approach does not scale efficiently when the number of relevant characteristics increases. In this case, an individual-based model (IBM) may be more appropriate. Within this approach, mathematical equations for the average population are replaced by individual rules of change that will be solved by computer simulation (36,(41).

4.1.2 Stochastic models

Most epidemic compartmental models are solved in a deterministic way; that is, predictions are determined entirely by their initial conditions, the set of underlying equations, and the assumed values of the parameters. However, when the number of infectious individuals is small, the homogeneous mixing assumption may not be appropriate. Furthermore, the stochastic nature of transmission and recovery events are critically important to understanding disease dynamics. This motivates the use of stochastic models, where S, I, and R as considered to be random variables.

One example is the Reed–Frost model, a simple stochastic epidemic mathematical model where population is divided into S and I. In the discrete-time version, let I_t be the number of individuals who got infected at time t and p the probability that a person encounters another person in one time-step and the contact results in disease transmission. A susceptible individual escapes infection at time t+1 if he/she avoids getting infected from the infected people at time t, so this happens with probability $(1-p)^{I_t}$.The probability to get infected is then given by $1-\ (1-p)^{I_t}$. Individual's classification may be updated only when time changes from t to t+1. The model can be written as:

$$I_{t+1} \sim Bin(S_t, 1-(1-p)^{I_t})$$

$$S_{t+1} = S_t - I_{t+1}$$

Overall, the Reed-Frost model remains a valuable tool for understanding the basic principles of infectious disease spread in small populations. To improve its limitations, some extensions have been proposed (42,43). Other stochastic methods are explained in (44).

4.2 Statistical-based models

Multitude of different statistical models have been widely used in epidemiology for the analysis of infectious diseases, epidemics, and vaccination data. Generalized linear models, mainly Poisson or negative binomial regression models, are commonly used to describe disease counts and rates (45,46). Here, the variability in the response (dependent) variable Y is explained by means of one or more of independent or control variables X. The conditional mean μ of the distribution depends on the independent variables as follows:

$$E(Y|X) = \mu = g^{-1}(X\beta),$$

where E(Y | X) is the expected value of Y conditional on X; Xβ is the linear predictor, a linear combination of unknown parameters β; and g is the link function. If Y follows a Poisson or negative binomial distribution, the link function is the log link and so:

$$\mu = exp(X\beta).$$

The unknown parameters β are typically estimated with maximum likelihood, maximum quasi-likelihood, or from a Bayesian viewpoint (47).

Poisson regression models have been the basis for the implementation of other more sophisticated models to describe epidemics. Held et al. (48) proposed a two-stage model where the counts of disease follow a Poisson or negative binomial observation model with two components: 1) a parameter-driven component that relates disease incidence to latent parameters describing endemic seasonal patterns, which are typical for infectious disease surveillance data, and 2) an observation-driven or epidemic component, which is modelled with an autoregression on the number of cases at previous time points and explains the increase in disease incidence due to epidemics. Let Y_t be the number of diseases counts at time t (t=1, …, n), following a Poisson distribution. The first model proposed by Held et al. is formulated as:

$$E(Y_t|Y_{t-1}) = \mu_t = (\nu + \lambda Y_{t-1})$$

Where $\nu > 0$ is the endemic component and λy_{t-1} is the epidemic one, with $0 < \lambda < 1$. Under certain conditions, it is interpreted as the R_0 (49).

A statistical discrete-time approach to conventional compartmental models was developed in Corberán et al. (50). The proposed model can be viewed as a flexible stochastic SIR model, which is analysed from a Bayesian viewpoint. Similar to Held et al., the authors model Y_t with a binomial distribution with parameters (S_{t-1}, p_t), where the probability of infection at time t, p_t, depends on the number of previous infections and is influenced by a transmission rate. Seasonality is accounted for through a Fourier series (50). These two methodologies will be explained in more depth in Chapter two, as they serve as the motivation for our proposed model in that chapter.

Held et al. endemic-epidemic model has also been expanded to accommodate either spatiotemporal data or age-structured scenarios for the modelling of multivariate time series. Within the scenario of infectious disease counts in neighbouring small areas, let Y_{it} be the count of disease in area i, i=1, …, I. Then, Y_{it} is assumed to follow a negative binomial distribution with mean given by:

$$\mu_{it} = \nu_{it} + \lambda Y_{i,t-1} + \phi \sum_{j \neq i} w_{ji} Y_{j,t-1},$$

where ν_{it} denotes the endemic component, λ captures the epidemic influence from the previous time step , and $\phi \sum_{j \neq i} w_{ji} Y_{j,t-1}$ incorporates spatial dependencies with weights w_{ji}. This model is designed to handle the intricacies of multivariate time series modelling, considering both the temporal and spatial dimensions of infectious disease counts (51).

One of the most popular models to describe spatial dependencies in the analysis of spatial health-related data is the Besag-York-Mollié (BYM) model, which incorporates random effects to capture unstructured and spatially structured heterogeneity. Let us assume that the observed count of disease in area i, Y_i, follows a Poisson distribution:

$$y_i \sim Poisson(e_i \theta_i),$$

where e_i is the known expected number of cases in the area i, and θ_i is the relative risk in area i. Then, this relative risk is modelled as

$$log(\theta_i) = \mu + s_i + h_j,$$

where μ is the overall risk level for the study region. The spatially correlated random effect is represented by s_i , and it is modelled as an autoregressive CAR-normal:

$$s_i \sim \mathrm{N}(\bar{s}_i, \tau_i^2)$$

$$\bar{s}_i = \frac{1}{\sum_j \omega_{ij}} \sum_j \mathrm{s_j}\omega_{ij}$$

$$\tau_i^2 = \frac{\tau_s^2}{\sum_j \omega_{ij}}$$

The ω_{ij} define the relationship between areas i and j (ω_{ij}=1 if i and j are neighbours and 0 if they are not). The random effect h_j is the uncorrelated heterogeneity and follows a $N(0, \tau_h^2)$. The precision parameters τ_s^2 and τ_h^2 control the amount of variability of the random effects (52).

A novel Bayesian hierarchical model that simultaneously allows for risk estimation and cluster detection in a purely spatial context is presented in (53). The model assumes that there exists an unknown number of underlying risk levels that describe the risk surface. Small area count data are assigned to a risk class using independent and identically distributed Multinomial indicator vectors. Consequently, geographically separated small areas with similar risks can be grouped into the same class. Unlike previous model formulations, no spatial correlation is imposed, providing so a flexible methodology to describe different risk structures. We delve into this modelling chapter four.

Chapter 2

A multivariate age-structured stochastic model with immunization strategies to describe bronchiolitis dynamics.

Abstract

Bronchiolitis has a high morbidity in children under 2 years old. Respiratory syncytial virus (RSV) is the most common pathogen causing the disease. At present there is only a costly humanized monoclonal RSV-specific antibody to prevent RSV. However, different immunization strategies are being developed. Hence, evaluation and comparison of their impact is important for policymakers. The analysis of the disease with a Bayesian stochastic compartmental model provided an improved and more natural description of its dynamics. However, the consideration of different age groups is still needed, since disease transmission greatly varies with age. In this work, we propose a multivariate age-structured stochastic model to understand bronchiolitis dynamics in children younger than 2 years of age considering high-quality data from the Valencia health system integrated database. Our modelling approach combines ideas from compartmental models and Bayesian hierarchical Poisson models in a novel way. Finally, we develop an extension of the model that simulates the effect of potential newborn immunization scenarios on the burden of disease. We provide an app tool that estimates the expected reduction in bronchiolitis episodes for a range of different values of the uptake and effectiveness.

1. Background

Bronchiolitis is a common lower respiratory tract infection (LRTI) that mainly affects children under 2 years old, with the greatest burden occurring in infants younger than 6 months (54). Respiratory syncytial virus (RSV) is by far the most common cause of viral bronchiolitis (55). Previous data from the Valencia region of Spain showed that approximately 2 out of 10 children younger than 2 years of age are diagnosed of bronchiolitis, 3 out of 100 are hospitalized and 1.6 out of 100 are hospitalized with RSV bronchiolitis (56,57).

At present, palivizumab is the only prophylactic therapy available for RSV. This is a costly humanized monoclonal RSV-specific antibody that is given monthly to infants at increased risk of severe RSV infection: premature babies, infants with chronic lung disease and infants with congenital heart disease (58,59). However, different RSV immunization strategies, including maternal immunization, immunoprophylaxis with new monoclonal antibodies (mAbs) and paediatric immunization, are currently under development and could be available soon (60,61).

Immaturity of the infant's immune system does not allow active immunization until 2 months of age. However, this population subgroup is more susceptible to severe RSV infection. Therefore, passive immunization practices (maternal vaccination and mAbs administered at birth) have the potential to be the most effective strategies to protect individuals at an early age. ResVax is the most advanced maternal vaccine in clinical development. In a phase III multi-country, randomized, placebo-controlled trial evaluating its efficacy against RSV-LRTI in infants from birth to 90–180 days of life, ResVax showed 44% efficacy in reducing RSV-LRTI hospitalization. However, it failed to meet the primary outcome of prevention of medically significant LRTI (62). A phase III trial showed the potential of mAb nirsevimab to protect infants for an entire RSV season with just one injection. A single dose of this mAb with an extended half-life resulted in a lower incidence of medically attended RSV-associated LRTI and hospitalizations than with a placebo in healthy preterm infants entering their first RSV season. They were approximately 70% and 80% lower respectively (63).

Before the licensure of these products for RSV and the implementation of a particular preventive strategy, policymakers will need to assess their potential health and economic benefits. This requires an exhaustive evaluation of the real burden of disease and the simulation of disease incidence under different immunization scenarios.

Different studies have been proposed to analyse either bronchiolitis or RSV dynamics. Most of these studies consider classical regression models to describe causal relationships between risk factors (baseline characteristics or clinical features) and disease counts (57,64–68). Deterministic compartmental models based on differential equations have also been commonly implemented to model RSV dynamics (69–72). Deterministic models assume that every individual has an equal probability of contacting every other

individual in the population, and so they may fail if the sizes of the compartments are not large enough to ensure homogeneous mixing. The incorporation of stochasticity into epidemic models is necessary for considering heterogeneous mixing of individuals in the population and also the inherent stochasticity in the transmission of disease (39,73–75). A Bayesian stochastic compartmental model in discrete time has improved the description of RSV disease (50).

Alternative stochastic approaches for the analysis of time series of infectious disease counts have been described by Held et al (48). They explain disease counts by means of a Poisson (or negative binomial) model with two components: a parameter-driven component that describes endemic seasonal patterns and an observation-driven component, which is an autoregression on past counts, that explains localized epidemics. This model can be extended in a multivariate setting to analyse time series that correspond to disease counts in different spatial units or different age groups jointly (49).

Systematic reviews of model-based evaluations of immunization strategies against RSV have been presented in (76,77). Many of the studies use mathematical models based on ordinary differential equations (see, for instance, (70,78,79)). Bayesian stochastic compartmental models have also been extended to simulate and assess the effect of a vaccination strategy that consists of vaccinating a proportion of newborns (80). Since clinical data are not yet available, models exploring the epidemiological and economic outcomes of potential RSV interventions adopt different assumptions regarding the target population, uptake rates, effectiveness, and duration of immunity. Therefore, the estimates that are obtained vary considerably.

In this chapter, we present a Bayesian multivariate age-structured model in discrete time to describe bronchiolitis dynamics in the region of Valencia in a population of children under 2 years of age. In particular, the population has been divided into four age groups and new infections are described considering interaction among these groups. We base our formulation on a Poisson model (or a negative binomial model if the data exhibit overdispersion). The main innovation is that the mean of the distribution at each time point can be seen as an observation-driven component where the autoregressive parameter varies stochastically over time. A parameter allowing for heterogeneous mixing of individuals in the population is also introduced to capture different contact

patterns. This modelling approach does not require information on susceptibles and it also provides a simplified framework to describe disease counts, since it avoids the modelling of complex transitions between the different compartments. To estimate the impact of potential prevention strategies for RSV, an extension of the model including simulation of bronchiolitis episodes under different passive immunization scenarios is considered.

2. The data

2.1 Population of interest

The region of Valencia, one of the 17 autonomous regions of Spain, has approximately 4,900,000 inhabitants. Around 2% of the Valencia population is younger than 2 years old (approximately 100,000 children). The Regional Health System (RHS) is divided into 241 health care districts structured into 24 health departments. It includes 34 public hospitals, 24 of them attending acute paediatric patients. This study includes children under the age of 2 years that were born in the region of Valencia between January 2009 and December 2012 that are covered by the RHS.

2.2 Data Sources

The region of Valencia has a health system integrated database (VID) that gathers health and sociodemographic data from 98% of the population (24). In particular, we used the population information system (an administrative database that collects and updates sociodemographic data from both residents and non-residents with access to public health services) to determine the population of interest. Primary care electronic medical notes (SIA) were implemented in 2006 and all medical visits are registered, and ICD coded. Hospitalizations were obtained using discharge reports from the Spanish Minimum Basic Data Set (MBDS).

2.3 Age-structured bronchiolitis cases

Because around 90% of bronchiolitis cases on children less than 2 years of age are treated in primary care offices and very few cases have a RSV microbiological confirmation, the analysis of laboratory-confirmed cases of RSV bronchiolitis is less useful for incidence estimation due to underreporting. Hence, it is more practical to analyse bronchiolitis-

associated outpatient visits and hospitalizations to evaluate the potential impact of RSV immunization strategies.

We analyse bronchiolitis episodes identified from hospitalization and primary care attendance through a search of the first appearance of the following ICD-9-codes: 466.1, 466.11 and 466.19, in MBDS and SIA. Based on McConnochie criterion, only the first health care encounter (either outpatient visit or hospitalization) with a bronchiolitis ICD-9-MC code was counted as a case of bronchiolitis.

Considering that children aged less than 6 months have an elevated morbidity, we consider four age groups: 0-5, 6-11, 12-17, and 18-23 months of age. This grouping allows us to analyse in greater detail both transmission dynamics and the impact of potential immunization measures. Weekly aggregated cases from July 2010 (week 06/28 – 07/04) to December 2012 (week 12/24 – 12/30) are considered (T = 131 weeks in total). Since this is a population-based cohort study, data for all the age groups were not available until July 2010, and so data from January 2009 to June 2010 are excluded in the analysis. In total, the study includes 30,555 cases of bronchiolitis, 14,635 (47.90%) of them are in children aged less than 6 months, 10,600 (34.69%) cases in age group 6-11 months, 3,462 (11.33%) in age group 12-17 months, and 1,858 (6.08%) in age group 18-23 months. Figure 2.1 shows the time plot of the series for the four age groups. As can be seen, there is a higher incidence for younger children and a clear seasonal pattern, with more cases during the weeks of autumn and winter. These data can be downloaded from https://rotapp.shinyapps.io/APP_IMPACTO_RSV.

3. Model description.

3.1 Background

The most commonly used models in the analysis of infectious disease counts are compartmental models, that divide the population being studied into different compartments according to disease status and describe the evolution of infection through changes in the number of individuals in each compartment. Corberán et al. (50) proposed a Bayesian stochastic compartmental model in discrete time to describe RSV dynamics in the Region of Valencia. Let $y_t, t = 1,2, \dots, T,$ be the number of newly infected individuals at week t. The model assumes that:

$$y_t \sim Bi(S_{t-1}, p_t), \tag{2.1}$$

S_{t-1} being the susceptible population at time $t-1$, which is weekly updated using some recursion equations, and p_t the probability of becoming infected at time t. To take into account the transmissible nature of the infection, p_t is modelled as:

$$p_t = \frac{y_{t-1}^h \cdot \exp(r_t)}{1+y_{t-1}^h \cdot \exp(r_t)}, \tag{2.2}$$

where the mixing parameter h allows for heterogeneous mixing of individuals in the population and $r_t = \alpha_0 + \sum_{w=1}^{W} \left(\beta_w \cdot \sin\left(\frac{2\cdot\pi\cdot w\cdot t}{52}\right) + \gamma_w \cdot \cos\left(\frac{2\cdot\pi\cdot w\cdot t}{52}\right)\right) + \varepsilon_t$. By allowing the transmission rate $\exp(r_t)$ to vary over time, the stochastic model provided an improved and accurate description of the pattern of disease.

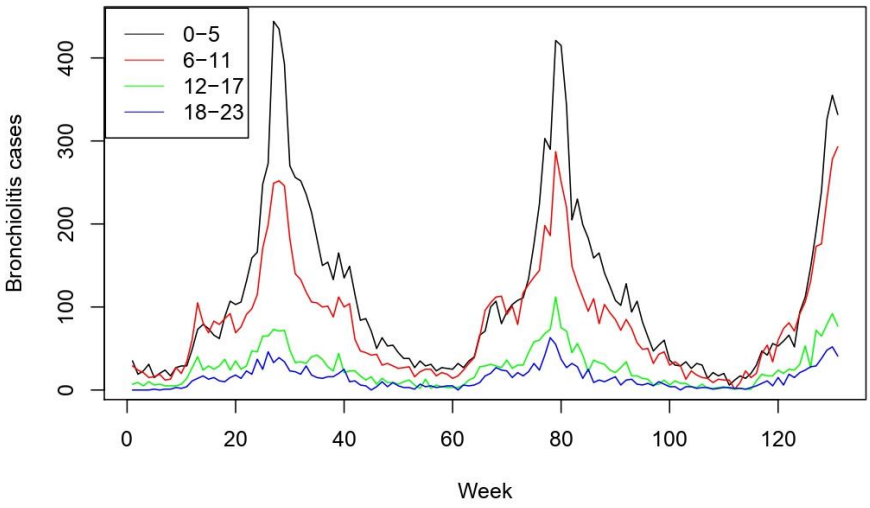


Figure 2.1: Weekly counts of bronchiolitis for the four age groups from July 2010 (week 06/28 – 07/04) to December 2012 (week 12/24 – 12/30).

When the focus is on describing counts of new infections, the use of a Bayesian hierarchical model provides an alternative framework. Held et al. (48) proposed an stochastic model for the analysis of disease counts based on a Poisson (or negative binomial) model with two components, which describe endemic seasonal patterns and localized epidemics. The starting point was a simple branching process model, which was latter extended to include seasonal terms in the endemic rate or to adjust for overdispersion. In particular, weekly new counts of disease $y_t, t = 1,2,\ldots,T$, are modelled as:

$$y_t \sim \ Poisson(\nu_t + \lambda_t \cdot y_{t-1}), \quad (2.3)$$

where ν_t is the parameter-driven (or endemic) component and $\lambda_t \cdot y_{t-1}$ the observation-driven (or epidemic) component, which allows for occasional outbreaks. A main feature of the proposed model is that the autoregressive parameter λ is allowed to vary over time. To do that, a Bayesian changepoint model with unknown number of changepoints is used to capture sudden changes in infectiousness. The endemic component is described by:

$$\nu_t = \exp\left(\alpha_0 + \beta \cdot \sin\left(\frac{2\cdot\pi\cdot t}{52}\right) + \gamma \cdot \cos\left(\frac{2\cdot\pi\cdot t}{52}\right)\right). \quad (2.4)$$

Multivariate extensions of that model can be found in (81) for the joint analysis of multiple time series of counts, where each component corresponds to a geographical region or a certain age group. Let y_t^j denote the count of disease in age group j at week t. In the multivariate scenario, counts of disease can be described as:

$$y_t^j \sim \ Poisson(\nu_t^j + \lambda \cdot y_{t-1}^j),$$

$$\nu_t^j = \exp\left(\alpha_0^j + \alpha_1 \cdot t + \sum_{w=1}^{W}\left(\beta_w \cdot \sin\left(\frac{2 \cdot \pi \cdot w \cdot t}{52}\right) + \gamma_w \cdot \cos\left(\frac{2 \cdot \pi \cdot w \cdot t}{52}\right)\right)\right),$$

where parameters α_0^j allow for different incidence levels in the different age groups. The epidemic component here depends only on previous counts in the corresponding age group.

A more general model considering previous counts in other age groups as potential explanatory variables is formulated as:

$$y_t^j \sim \ Poisson\left(\nu_t^j + \lambda \cdot y_{t-1}^j + \phi \cdot \sum_{i \neq j} y_{t-1}^i\right),$$

where the additional parameter ϕ captures the autoregressive effect of the other age groups. Paul et al. (82) extended this multivariate model by allowing the autoregressive parameters to depend on the age-group; that is, λ^j and ϕ^j for each time series. Yet, in these multivariate models, the autoregressive parameters λ and ϕ are not allowed to vary over time.

3.2 Our proposal

Let y_t^j denote the count of disease in age group $j, j = 1,2,3,4$, observed at week t, $t = 1,2,\dots,T = 131$. We use a Bayesian hierarchical model, which provides a

straightforward framework to describe counts of new infections at each week. Our model assumes that the probability governing the counts is a Poisson distribution or a negative binomial distribution if the data exhibit overdispersion. In those situations, in which the population size is not large enough in comparison with the observed counts of disease, the binomial distribution could be used in a similar way.

Seasonal variation in disease transmission that is persistent with a stable pattern is modelled through sine-cosine waves which are common for all groups. This is a sensible assumption since annual epidemic peaks occur simultaneously in the different age groups. On the other hand, it is reasonable to assume that contacts among the different age groups considered here are equally likely to occur. Hence, counts of disease at time $t-1$ are summed up to describe new infections. Also, heterogeneous mixing of individuals in the population is allowed. Finally, due to the epidemiology of bronchiolitis, it is important to consider a different transmission rate for each age group.

Taking into account these considerations, the proposed model is given by the following equations:

$$y_t^j \sim Poisson(\mu_t^j), \tag{2.5}$$

$$\mu_t^j = \exp(\lambda_t^j) \cdot \left(\sum_{i=1}^{4} y_{t-1}^i\right)^h, \tag{2.6}$$

$$\lambda_t^j = \alpha_0^j + \beta \cdot \sin\left(\frac{2 \cdot \pi \cdot t}{52}\right) + \gamma \cdot \cos\left(\frac{2 \cdot \pi \cdot t}{52}\right) + \varepsilon_t, \tag{2.7}$$

where the autoregressive parameter $\exp(\lambda_t^j)$ is allowed to vary stochastically over time by means of random effects $\{\varepsilon_t \sim N(0, \sigma_\varepsilon^2)\}$ that represent unspecified features of week t. To avoid overfitting, these random effects are common for all the age groups; that is, if counts of disease at a particular week are higher (or lower) than expected for some age groups, that increase (or decrease) should also be observed in the remaining age groups. $h \in [0,1]$ is the mixing parameter ($h = 1$ would correspond to the assumption of mass action). Non-informative flat prior distributions are considered for parameters α_0^j, β, and γ. The uniform distribution in the interval [0,1] is considered as a prior for h and the uniform distribution in the interval (0,2) is considered for the standard deviation σ_ε.

To adjust for overdispersion, the negative binomial distribution can be used instead. In that case, the model can be formulated as:

$$y_t^j \sim NegBin(p_t^j, k), \tag{2.8}$$

$$p_t^j = \frac{k}{k+\mu_t^j}, \tag{2.9}$$

where μ_t^j is defined as in Equation (6). Hence, the mean of the distribution is equal to μ_t^j and the variance is given by $\mu_t^j + \frac{(\mu_t^j)^2}{k}$. The second parameter $k > 0$ incorporates the extra-Poisson variation (the limiting case $k = \infty$ corresponds to the Poisson distribution). As a prior distribution for parameter k, we consider the Gamma distribution $Ga(1,0.01)$. It is important to emphasize that our formulation can be seen as a Poisson model with only one component, the observation-driven component (48,81), modified in a novel way to allow for contact heterogeneity. By considering a time-varying autoregressive parameter, the model properly describes both endemic an epidemic period. As shown in the next section, this formulation will allow us to simulate disease incidence in a scenario with immunization measures. On the other hand, the mean of the distribution is equivalent to that proposed in (50), where the process intensity mean was dependent on the product of S_{t-1}, y_{t-1}^h and the transmission rate $\exp(r_t)$. In our modelling framework, where information on susceptibles is not incorporated, the autoregressive parameter $\exp(\lambda_t^j)$ describes how infected individuals in the previous week produce new infections.

3.3 Our extension with immunization strategies

We show here how the proposed model can be used to simulate the impact of passive immunization through maternal vaccination or mAb administered at birth. The main parameters to consider when estimating the effects of a particular immunization strategy are the uptake, the efficacy and the duration of protection.

Based on vaccination coverage data provided by the Ministry of Health, Consumer Affairs and Social Welfare of Spain (https://www.mscbs.gob.es/profesionales/saludPublica/prevPromocion/vacunaciones/calendario-y-coberturas/coberturas/home.htm) and the clinical trial results presented in (62) and (63), the following assumptions are incorporated into the simulation process:

In the case of maternal vaccination:

(a) A range of vaccine uptake from 50 to 85%.

(b) 40% efficacy in reducing RSV bronchiolitis episodes.

(c) Duration of protection of 6 months.

With respect to the use of mAb:

(a) The uptake parameter can be considered to be 95%.

(b) 70% efficacy in reducing RSV bronchiolitis episodes.

(c) 6 months of induced immunity.

Because the duration of protection is assumed to be 6 months, children are immune from birth to 6 months, when they move to the next age group. Hence, in our modelling scheme, the evaluation of these two strategies come down to a decrease in the number of cases in the first age group (0-5 months) throughout the simulation study. Since we analyse total counts of bronchiolitis, it is also important to take into account that approximately 70% of bronchiolitis cases are due to RSV.

Table 2.1 shows the percentage of children in the age group 0-5 months that should be removed according to the specified values for the uptake and the efficacy parameters. These values are also in accordance with the assumptions made in some of the studies discussed in (76).

Maternal vaccination			**mAb at birth**		
Uptake	Efficacy	% removed	Uptake	Efficacy	% removed
50%	40%	20%	95%	70%	66.50%
85%	40%	34%			

Table 2.1: Possible scenarios in the evaluation of immunization strategies.

Let π_{rsv} represent the proportion of bronchiolitis cases due to RSV and π_{im} the proportion of children in the age group 0-5 months that are immune as a result of the implementation of an immunization strategy. For age group 1 (0-5 months), the number of infected children at week t can be estimated from our model as:

$$\hat{y}_t^1 = (1-\pi_{im})\cdot\pi_{rsv}\cdot\left(\exp(\lambda_t^1)\cdot\left(\sum_{i=1}^4 \hat{y}_{t-1}^i\right)^h\right) + (1-\pi_{rsv})\cdot\left(\exp(\lambda_t^1)\cdot\left(\sum_{i=1}^4 \hat{y}_{t-1}^i\right)^h\right); \tag{2.10}$$

that is, a proportion π_{im} is removed from the number of RSV bronchiolitis cases that would be expected from contacts with infected children at week $t-1$. Note that, under the immunization scenario, contacts of immune children will not end in contagion.

For the age groups $j = 2,3,4$, infected children at week t can be estimated as:

$$\hat{y}_t^j = \exp(\lambda_t^j)\cdot\left(\sum_{i=1}^4 \hat{y}_{t-1}^i\right)^h. \tag{2.11}$$

In the previous equations, $\exp(\lambda_t^j)\cdot\left(\sum_{i=1}^4 \hat{y}_{t-1}^i\right)^h$ represents the mean μ_t^j of the distribution governing the counts, which is either the Poisson or the negative binomial distribution, based on the simulated counts under the immunization scenario. Parameters

$\{\exp(\lambda_t^j)\}$ and h are, respectively, the autoregressive parameters and the mixing parameter of the model without an immunization strategy (see Equation (2.6)). It is important to emphasize that these parameters represent features of bronchiolitis dynamics (how the disease spreads at each time point for the different age groups) and contact patterns that do not depend on the number of infected children at previous weeks. Hence, they can be used to simulate the number of infections after a particular preventive intervention has been implemented.

4. Results

R Statistical Software (Foundation for Statistical Computing, Vienna, Austria) and WinBUGS program (83) were used to perform the analysis using MCMC simulation methods. A total of 25,000 iterations were used as the burn-in period of the MCMC. After that, 75,000 iterations were run and only 1 in every 150 of them was kept to reduce correlation. 2 chains were simulated, and so $M = 1{,}000$ values were simulated in total from the posterior distribution. MCMC convergence was assessed by visual inspection of the trace plots, the Brooks-Gelman-Rubin scale reduction factor (Rhat equal to 1 means good convergence) and the effective sample size (n.eff above 100 means good convergence). All statistical analyses of the study are completely reproducible. The BUGS code used can be found as supplementary material to the paper.

4.1 Results without immunization strategies

We fitted both the Poisson and the binomial negative model to the bronchiolitis data. The corresponding DIC (84) values are, respectively, 3811.77 (pD 105.89) and 3710.97 (pD 83.96), so adjusting for overdispersion provides a better fit. The posterior mean of parameter k is 44.45, being the 95% credible interval [32.65,59.41]. Figure 2.2 displays the posterior mean bronchiolitis temporal profile (red line) together with real counts of disease (black line). As can be seen, the proposed model is able to accurately recover disease dynamics in all the age groups.

Figure 2.3 shows the estimated autoregressive parameter $\exp(\lambda_t^j)$ in each age group and time point together with its seasonal component, which is given by:

$$\exp\left(\alpha_0^j + \beta \cdot \sin\left(\frac{2 \cdot \pi \cdot t}{52}\right) + \gamma \cdot \cos\left(\frac{2 \cdot \pi \cdot t}{52}\right)\right).$$

As expected, previous counts of disease have a higher impact on youngest children. In fact, the transmission of disease decreases as the age increases. This figure also demonstrates that seasonality has a strong influence on transmission dynamics. However, because the seasonal pattern varies slightly from year to year, the incorporation of random effects accounting for stochasticity in the transmission is fundamental to provide a more accurate description of the data. These results agree with those obtained in (50). In that paper, the proposed model with a stochastic transmission rate led to an improved goodness of fit in comparison with its counterpart where the seasonal pattern repeated over time; that is, there were no weekly random effects in the model assumed for the transmission rate.

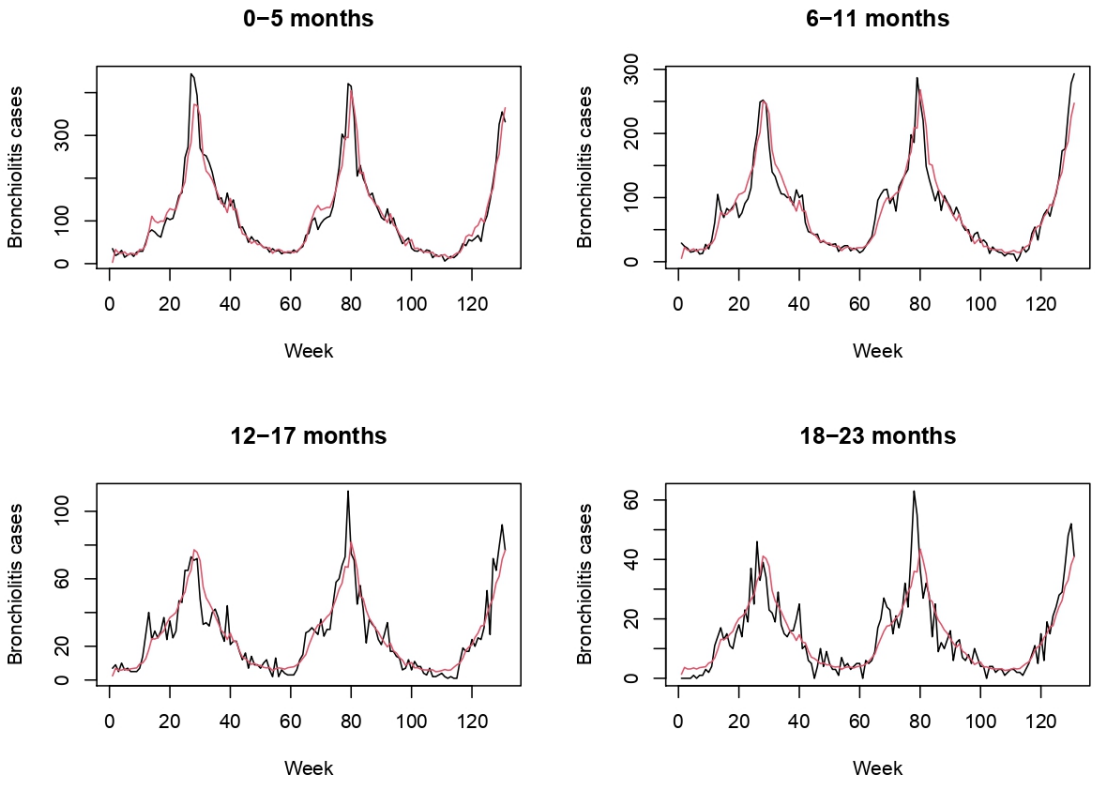


Figure 2.2: Model accuracy. Bronchiolitis counts (black line) together with posterior mean estimates (red line) from July 2010 to December 2012 (T = 131 weeks) obtained with the proposed model based on the negative binomial distribution.

4.1.1 Model comparison.

In order to assess the performance of our model, we also analysed the bronchiolitis data with a Poisson model whose mean is given by the sum of an endemic and an epidemic components. Taking into account the assumptions made in this particular case study (common seasonality for all ages and contacts among the different age groups that are equally likely to occur), the Poisson model with two components is formulated as:

$$y_t^j \sim Poisson\left(\nu_t^j + \lambda^j \cdot \sum_{i=1}^{4} y_{t-1}^i\right) \tag{2.12}$$

$$\nu_t^j = \exp\left(\alpha_0^j + \beta \cdot \sin\left(\frac{2\cdot\pi\cdot t}{52}\right) + \gamma \cdot \cos\left(\frac{2\cdot\pi\cdot t}{52}\right)\right). \tag{2.13}$$

To ensure that the mean of the Poisson is non-negative, parameter λ^j is modeled as $\lambda^j = \exp(\delta^j)$, with $\delta^j \sim N(0, \sigma_\delta)$.

Figure 2.4 displays the posterior mean bronchiolitis temporal profile (red line) corresponding to this Poisson model with two components together with real counts of disease (black line). The DIC value for this model is 4506.18 (pD 9.88).

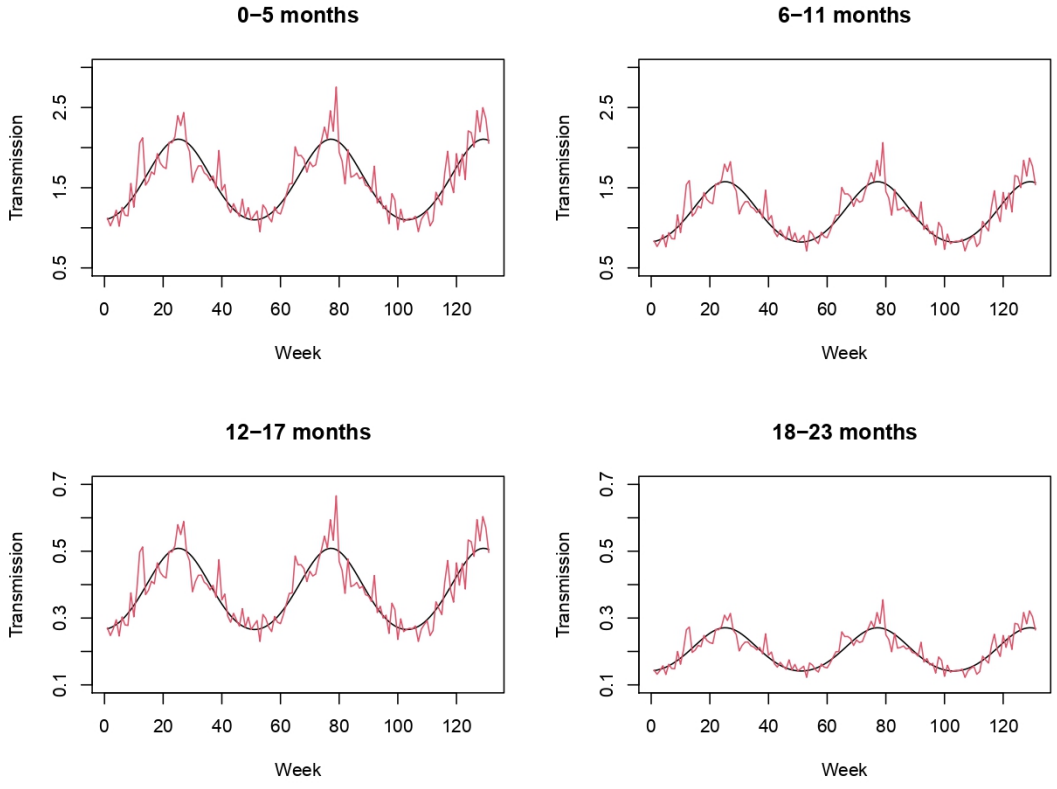


Figure 2.3: Estimated autoregressive parameter (red line) together with its seasonal component (black line).

This two-component model is able to describe the overall dynamics of bronchiolitis. However, the Poisson model proposed here provides a better fit, as judged by a lower DIC (3811.77 against 4506.18). This may be since the epidemic component, which depends on a time-constant autoregressive parameter, cannot explained the stochasticity inherent in transmission dynamics.

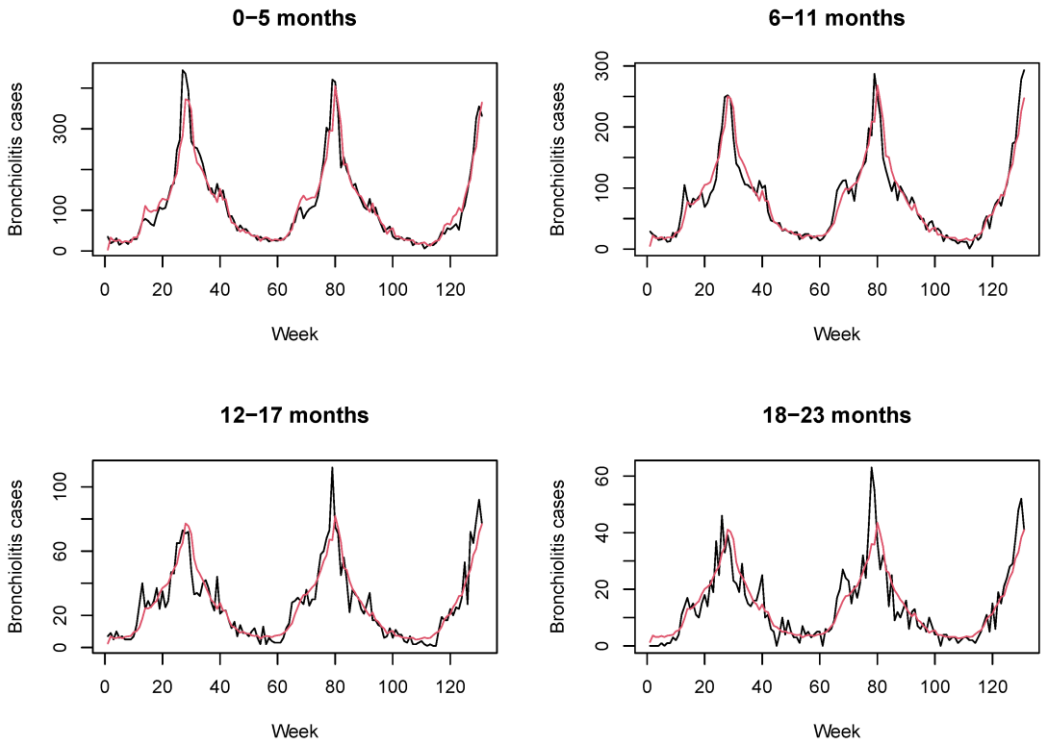


Figure 2.4: Bronchiolitis counts (black line) together with posterior mean estimates (red line) obtained with a Poisson model with two components.

4.2 Results with immunization strategies

Once we have obtained a sample from the posterior distribution of the parameters of our model without an immunization program, we can evaluate the effect of the immunization strategies previously described. We assume here that the immunization strategy started six months before the first week of July 2010 (week 06/28 – 07/04), so that the specified percentage (see Table 2.1) of children in the age group 0-5 months can be assumed to be immune at time $t = 1$ of our analysis. We simulate the evolution of bronchiolitis counts using Equations (10) and (11), assuming $\pi_{rsv} = 0.7$. Note that for each iteration of the

MCMC simulation, we have one simulated value of the regression parameters and the mixing parameter: $\{\exp(\lambda_t^{j(m)})\}$ and $h^{(m)}$, $m = 1,2,\ldots,M = 1{,}000$, and so we have M simulated values for the number of infections for each age group and week. Using these values we can get point estimates (for instance, the mean of the simulated values) as well as credible intervals.

Figure 2.5 shows real counts of bronchiolitis and point estimates (mean values) for three simulation scenarios, corresponding to π_{im} = 0.2 (green line), 0.34 (blue line), and 0.67 (red line). As expected, the number of bronchiolitis cases decreases as the proportion of immune children in the age group 0–5-month increases. Because of herd protection, the decrease can be observed in all the age groups. The decrease expected in each age group and the total decrease are shown in Table 2.2.

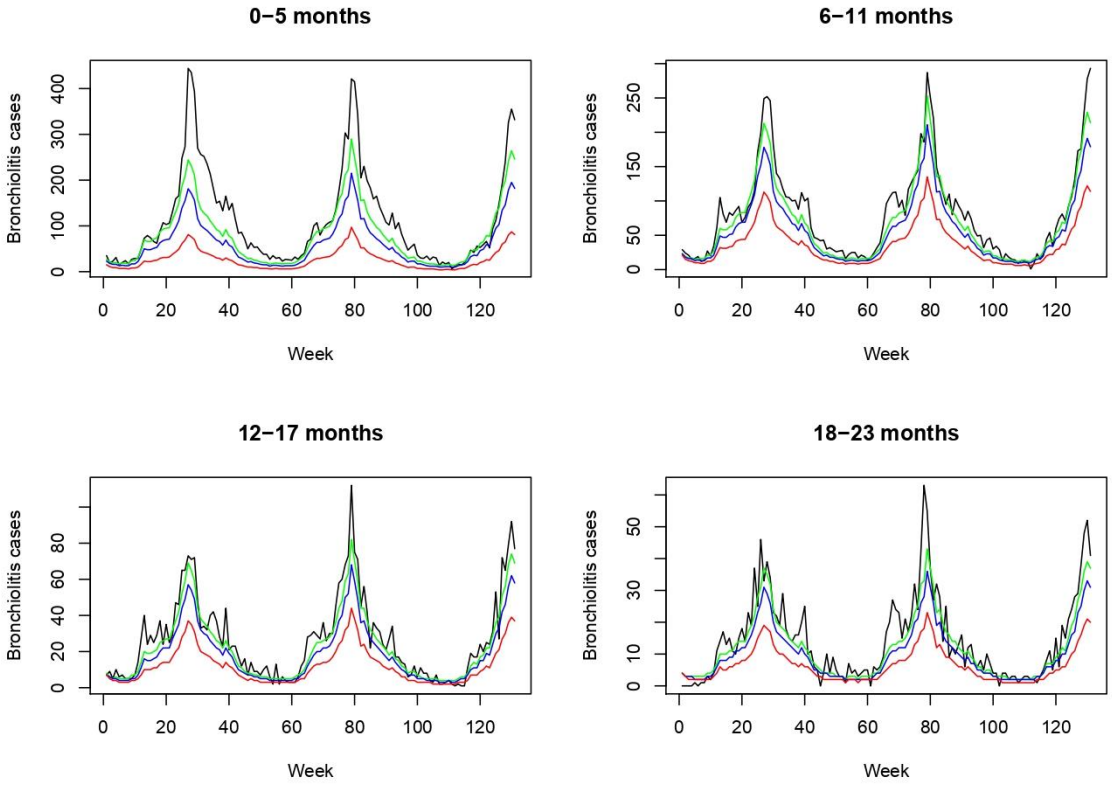


Figure 2.5: Bronchiolitis counts (black line) together with simulated counts (mean values) for three different immunization proportions: π_{im} = 0.2 (green line), 0.34 (blue line), and 0.67 (red line).

π_{im}	0-5 months	6-11 months	12-17 months	18-23 months	Overall
0.2	28.94%	14.63%	15.48%	16.25%	21.68%
0.34	7.48%	28.75%	29.78%	30.03%	7.92%
0.67	6.32%	4.17%	54.56%	54.84%	64.87%

Table 2.2: Expected decrease in the number of bronchiolitis counts for three different immunization proportions: π_{im} = 0.2, 0.34, and 0.67.

We have also developed an R Shiny application (available at https://rotapp.shinyapps.io/APP_IMPACTO_RSV) that simulates the evolution of bronchiolitis that would be obtained by varying the uptake and the efficacy parameters, as well as the percentage of bronchiolitis cases due to RSV (more information in annex 2.2).

5. Conclusions

We have developed an age-structured stochastic model that allows us to accurately explain bronchiolitis dynamics in the region of Valencia. Our modelling approach combines ideas from compartmental models and Bayesian hierarchical Poisson models in a novel way. By using a hierarchical Poisson (or a binomial negative) model, we do not require information on susceptibles and we avoid the modelling of complex transitions between the multiple compartments. Unlike standard formulations, the main innovation of our model is that the mean of the distribution at each time point depends only on an observation-driven component where the autoregressive parameter is allowed to vary stochastically over time. A mixing parameter is also introduced to capture heterogeneous mixing of individuals in the population. By modelling the dynamics of disease as a function of previous counts, we can simulate disease incidence under different immunization scenarios. Note that the incorporation of a parameter-driven component would hinder simulation of disease evolution when a decrease in the number of cases at previous weeks has been obtained as a result of the assumed immunity.
Even though we have developed the model for the analysis of bronchiolitis, it can be adapted for other infectious diseases with (or without) a seasonal pattern, replacing the transmission rate according to the nature of disease.

The extension proposed in this paper provides a useful framework to address one of the important needs in RSV bronchiolitis incidence control: the implementation of an

immunization strategy. We have assessed here the effects of passive immunization through maternal vaccination or mAb administered at birth. In our modelling scheme, the evaluation of these two strategies come down to a decrease in the number of cases in the first age group (0-5 months) throughout the simulation study. We have simulated the evolution of bronchiolitis counts for different values of the uptake and efficacy. The duration of immunity has been assumed to be equal to 6 months. If immunity lasts more than 6 months, the model can be adapted so that a percentage of children in the age group 6-11 months is also removed, since children would move to this age group being immune for a certain number of weeks. On the other hand, if immunity lasts less than 6 months, the percentage of children removed in the age group 0-5 months should be reduced accordingly.

The results obtained show that newborn immunization contributes to a substantial decrease in the number of bronchiolitis infections. Because of herd immunity, this decrease is also observed in all the age groups considered. As expected, the decrease is greater when the uptake and/or the efficacy increase. As the profiles of the products under development become better defined, further alignment will be possible.

We have developed an R Shiny application that simulates the evolution of bronchiolitis episodes for different values of the uptake and the efficacy, as well as the percentage of bronchiolitis cases due to RSV. This application is an easy tool that allows users and decision makers to interact with the simulation analysis and improves visualization of the results.

We have not assessed here the impact of immunization of high-risk children, such as preterm infants or children with high-risk comorbid conditions. It may be important to consider a separate analysis for these groups most at risk.

Another fruitful area for further research would be the consideration of smaller age groups, so that different immunization strategies such as vaccination of infants older than 3 months could also be tested. Stratification by gestational age may also be important when evaluating maternal vaccination strategies since transfer of maternal antibodies for preterm infants may be incomplete.

The content of this chapter was published in Int J Environ Res Public Health. 2021 Jul 17;18(14):7607. doi: 10.3390/ijerph18147607.

Chapter 3

Potential impact of nirsevimab and bivalent maternal vaccine against RSV bronchiolitis in infants: A population-based modelling study.

Abstract

A new monoclonal antibody (nirsevimab; Beyfortus®) and a bivalent prefusion RSV vaccine (Abrysvo®) for maternal immunization have been approved recently. This is a modelling study to estimate the potential impact of different immunization programs with these products on RSV-bronchiolitis. Population-based real-world data from primary care and hospitalizations were considered. RSV bronchiolitis dynamics in absence of these immunization scenarios were explained by a multivariate age-structured Bayesian model. Then, the potential impact was simulated under different assumptions including the most recent clinical trial data. Differences in endpoints, populations, and timeframes between trials make the two products' efficacy difficult to compare. Our results suggest that each strategy would effectively reduce RSV-bronchiolitis. A seasonal with catch-up program with nirsevimab would prevent up to 9,210 RSV bronchiolitis and up to 1,755 RSV-hospitalizations per 100,000 infants-year. A year-round maternal immunization program would prevent up to 5,837 RSV bronchiolitis and up to 1,282 RSV-hospitalizations per 100,000 infants-year.

1. Background

Respiratory Syncytial Virus (RSV) infections are a significant threat to public health, particularly for vulnerable populations such as babies, older adults, and those with compromised immune systems (85). The virus is the most common cause of acute lower respiratory tract infections in infants and is responsible for many deaths in low- and middle-income countries within the first 6 months of life (86). Evidence also links RSV infection with pneumococcal infection, positioning children at higher risk of severe disease (87).

Bronchiolitis is the most common presentation in RSV-infected infants, leading to hospitalization in about 2-4% of infants under one year old. Approximately 20% of infants with RSV infection require outpatient medical attention (88–90). RSV accounts for about 60% of bronchiolitis hospitalizations (57,86,91) and over 50% of primary care cases (92).

Until now, the only available prophylactic therapy to prevent RSV disease was the RSV-specific monoclonal antibody (mAb) palivizumab. However, its high cost and requirement for monthly injections restrict its widespread use, making it recommended only for children with "high-risk" conditions like prematurity and bronchopulmonary disease (93).

Two recent prophylactic drugs have been approved: a new monoclonal antibody (nirsevimab; Beyfortus®; NmAb) for passive immunization that has shown a long-lasting (at least 150 days) protection against RSV disease in infants (94–96), and a bivalent prefusion RSV vaccine (Abrysvo®) for maternal immunization (MI) that has also demonstrated efficacy in preventing RSV in infants from birth up to six months of age (97).

Several modelling studies, including compartmental deterministic models, static cohort models, decision tree models, individual-level transmission models, and Bayesian models, have evaluated the impact of these new prophylactic interventions for RSV (77,80,98). They suggest that both the NmAb and MI would effectively reduce hospitalizations, with seasonal NmAb demonstrating a higher impact (99). MI has also shown potential significant reductions in mortality and morbidity in low- and middle-income countries (100). However, further research is needed to assess the realistic impact of these interventions in different populations, considering RSV burden in ambulatory care settings and incorporating recent clinical trial data to improve model assumptions (60,77).

Comparing the potential impact of the two interventions is challenging due to differences in the endpoints of the clinical trials, which can introduce bias depending on the chosen endpoint. To address this, our study simulates the potential impact of each intervention on the prevention of RSV bronchiolitis in infants under 12 months old within the same

population. Bronchiolitis serves as a proxy for RSV infections due to its well-modelled dynamics and its representation of a significant proportion of RSV infections in infants. We employ an enhanced and simplified model-based approach (98), that accurately describes RSV-bronchiolitis dynamics and serves as the basis for estimating the impact of each of the two new strategies, NmAb, and MI. The results of these estimates might be useful for public health decisions.

2. Methods

A multivariate age-structured model was developed to explain the dynamics of RSV bronchiolitis in children under 2 years old, using population-based bronchiolitis cases from the Valencia Integrated Databases (VID) in the Valencia Region of Spain (24). An extension of this model was implemented to simulate the impact of RSV NmAb when administered in infants or the bivalent maternal vaccine (MI) on RSV bronchiolitis in infants (101).

2.1 Study population and period

The study population included all children under 2 years old residing in the Valencia region of Spain from November 2010 to the end of January 2016, including 290,484 infants. Our model focuses on bronchiolitis cases within this age group to explain RSV bronchiolitis dynamics. However, in assessing the impact of immunization strategies, we only estimated the impact in infants up to 11 months of age.

2.2 Data sources

Data for this study were sourced from the Valencia Health System Integrated Database (VID) (24) which consists of multiple public, population-wide electronic databases, including the population-based administrative database (SIP) to determine the study population, Primary Care (PC) database (SIA), which contains medical information for each ICD9-coded PC visit, hospital discharge database (Minimum Basic Data Set; MBDS) which collects diagnoses and procedures at discharge (ICD9 until 2016 and ICD10 afterward), and the microbiological dataset (RedMIVA) used to identify RSV lab-confirmed hospitalizations.

2.3 Identification of RSV-Bronchiolitis cases

First, we identified the incidence of all-cause bronchiolitis by searching for the first occurrence of the following codes (ICD-9 466.11, 466.1, 466.19 or ICD10 J21.0, J21.8, J21.9) in PC (SIA) or hospital (MBDS) datasets. Then, the following two outcomes were studied: 1) RSV-bronchiolitis hospitalizations, defined as the first occurrence of the before-mentioned ICD-codes in any diagnostic position in hospital setting (MBDS) with a laboratory RSV-confirmation from RedMIVA; 2) RSV-bronchiolitis, defined as 54.9% of all-cause bronchiolitis (based on literature, 8) attended in PC (SIA) and RSV-bronchiolitis hospitalizations. We included both settings to be aligned with the definition used in clinical trials for RSV medically attended lower respiratory tract infections (RSV-MA-LRTI). When a PC bronchiolitis code was followed by a hospitalization within 30 days, it was assumed to be the same case and considered hospitalized bronchiolitis.

2.4 Age-structured time-series

Data were processed to obtain weekly age-structured time series, providing a detailed representation of RSV bronchiolitis cases over time (annex 3.2). Weekly counts of RSV-Bronchiolitis and RSV-bronchiolitis hospitalizations were considered for modelling. The data covered a total of 270 weeks, from November 2010 (week 48, 2010) to January 2016 (week 4, 2016). Cases were stratified into 13 age-groups, including monthly age-groups up to 12 months of age, and the 12-23 months group.

2.5 Model Description

Two multivariate age-structured Bayesian models (one for each outcome) were developed to explain the dynamics of the bronchiolitis in the absence of the two immunization scenarios. Both provided an accurate estimation of the real data in the different settings for all age-groups (annex 3.2). Then, an extension of this model was implemented to simulate the potential impact of each of the two RSV immunization programs (with NmAb or MI) depending on the immunization coverage. Models were based on a previously published model (chapter 2) (98), in which, a Bayesian Poisson (or negative binomial) model was developed to describe the disease dynamics when the data available are counts of new infections each week.

This study aimed to have a greater knowledge of the disease and more precise and realistic impact estimations considering the new results regarding efficacy and protection duration in clinical trials. Hence, we structured the first two years of life into smaller groups (13 age-groups instead four). Because we have smaller populations, we modelled counts of RSV bronchiolitis (and RSV bronchiolitis hospitalizations) with a Binomial distribution.

The model assumed that the distribution governing the counts was the Binomial distribution:

$$y_t^j \sim Bi\left(N_t^j, p_t^j\right)$$

Where, y_t^j is the number of potential RSV-bronchiolitis (or RSV-bronchiolitis hospitalizations), p_t^j, the probability of becoming infected and, N_t^j, was the population size in age-group j (j = 1, …, 13) at time t (t = 1, …, n=270).

The probability of becoming infected at time t for each age-group was modelled though the logit of p_t^j , $logit(p_t^j)$ as:

$$logit(p_t^j) = \log\left(\frac{p_t^j}{1-p_t^j}\right) = \lambda_t^j + h \cdot \log\left(\sum_{i=1}^{13} y_{t-1}^i\right) \quad (3.1)$$

This equation is equivalent to:

$$p_t^j = \frac{\exp\left(\lambda_t^j\right)\cdot\left(\sum_{i=1}^{13} y_{t-1}^i\right)^h}{1+\exp\left(\lambda_t^j\right)\cdot\left(\sum_{i=1}^{13} y_{t-1}^i\right)^h} \quad (3.2)$$

That is, the probability of infection p_t^j , depends on the previous number of infected children (in hospital and in PC) in all age-groups and a transmission rate $\exp\left(\lambda_t^j\right)$ that is allowed to vary stochastically over time. On the other hand, it is reasonable to assume that contacts among the different age groups considered here are likely to occur. Hence, counts of potential RSV-bronchiolitis at time t−1 are summed up to describe new infections. Parameter h allows for heterogeneous age-mixing of individuals in the population.

The transmission rate was modelled as:

$$\lambda_t^j = \alpha_0^j + \beta \cdot \sin\left(\frac{2\cdot\pi\cdot t}{52}\right) + \gamma \cdot \cos\left(\frac{2\cdot\pi\cdot t}{52}\right) + \varepsilon_t, \quad (3.3)$$

α_0^j, allows for different incidence levels in the different age groups. The simple harmonic wave was considered to account for seasonality (2). A common seasonality for all age

groups was considered. As shown in Supplemental material 1a, annual epidemic peaks occur simultaneously in all age-groups. $\varepsilon_t \sim N(0, \sigma^2_\varepsilon)$, is a stochastic random effects term which represents the unspecified features of week t. To avoid overfitting, these random effects are common for all the age groups; that is, if the probability of infection at a particular week are higher (or lower) than expected that increase (or decrease) should be also observed in all the age groups.

Non-informative flat prior distributions are considered for parameters α_0^j, β and γ. The uniform distribution in the interval [0, 1] is considered as a prior for the homogeneous mixing parameter h. The uniform distribution in the interval [0, 2] is considered for the standard deviation of the random errors.

R Statistical Software (Foundation for Statistical Computing, Vienna, Austria) and WinBUGS program were used to perform the analysis using MCMC simulation methods. A total of 25,000 iterations were used as the burn-in period of the MCMC. After that, 75,000 iterations were run, and only 1 in every 150 of them was kept reducing correlation. Two chains were simulated, so M=1000 values were simulated in total from the posterior distribution. MCMC convergence was assessed by visual inspection of the trace plots, the Brooks–Gelman–Rubin scale reduction factor (Rhat equal to 1 means good convergence) and the effective sample size (n.eff above 100 means good convergence). All statistical analyses of the study are completely reproducible.

2.6 Impact simulation assumptions

The impact of a program with NmAb or MI on RSV bronchiolitis was simulated under different assumptions including intervention efficacy and duration of protection, uptake rates, and immunization strategies.

2.6.1 Effectiveness and duration of protection

The assumed effectiveness of each product was based on reported clinical trial efficacy (95,97,102–104). Only direct effects were considered due to insufficient data on potential indirect effects.

For NmAb, pooled analyses of pivotal phase 2b and phase 3 studies showed a 5-month efficacy of 79.5% in preventing RSV-MA-LRTI and 77.3% efficacy for RSV-MA-LRTI

hospitalizations in infants during their first RSV season (104). NmAb appeared to have longer protection (95,105), f as according to the pharmacoeconomic model presented by the CDC (106), efficacy remained constant for the first 5 months, then linearly decreased to 0 by 10 months (Table 3.1).

For the MI, efficacy against hospitalization was 56% up to day 180, and 51.3% for RSV-MA-LRTI (97,102,107), then we assumed a , linear decrease to 0 by month 10 (Table 3.2). The clinical trial (108) described efficacy against RSV hospitalizations, which could potentially be higher for those with severe lower respiratory tract infections (LRTI), but this figure is not shown.

A sensitivity analysis of the effectiveness used in each immunization program was developed. We assumed 10% and 20% lower effectiveness than the main analysis for the first 5 months and 6 months for NmAb and MI, respectively. Then, a linear decay to 0 was used.

NmAb		
	RSV-bronchiolitis hospitalizations	**RSV-bronchiolitis**
Months of protection	Effectiveness	
0,1,2,3,4	77.3	79.5
Decay	*y=-12.88*x+128.83*	*y= -13.25*x+132.5*
5	66.3	64.4
6	53.0	51.6
7	39.8	38.7
8	26.5	25.8
9	13.3	12.9
10	0.0	0.0

Table 3.1: NmAb effectiveness values by month after administration, for RSV-bronchiolitis hospitalizations and RSV-bronchiolitis. Flat effectiveness was assumed for the first 5 months of protection, followed by a linear decay to 0. Equations represent the decay were, x, are the specific month after immunization and, y, the estimated effectiveness.

MI		
	RSV-bronchiolitis hospitalizations	**RSV-bronchiolitis**
Months of protection	Effectiveness	
0,1,2,3,4,5	56.8	51.3
Decay	*y=-11.36*x+113.6*	*y= -10.26*x+102.6*
6	45.4	41.0
7	34.1	30.8
8	22.7	20.5
9	11.4	10.3
10	0.0	0.0

Table 3.2: MI effectiveness values by month after administration, for RSV-bronchiolitis hospitalizations and RSV-bronchiolitis. Flat effectiveness was assumed for the first 6 months of protection, followed by a linear decay to 0. Equations represent the decay were, x, are the specific month after immunization and, y, the estimated effectiveness.

2.6.2 Immunization Programs

RSV season was considered between November-March (78). For NmAb, a seasonal program was considered, immunizing all newborns from November to March, with a catch-up immunization of infants born from April to October when they are under 7 months of age at the beginning of the RSV season.

For MI, two scenarios were considered: seasonal vaccination to prevent infections in children born right before or during the RSV season which would mean to immunize pregnant women when they expect to deliver between November and March, and year-round vaccination (Figure 3.1).

2.6.3 Immunization Coverage

As there might be different immunization coverages depending on the area, health system or program, we carried out a sensitivity analysis with different uptake levels (55%, 65%, 75%, 85% and 95%), so that the estimated impact of each program can be assessed.

Elegible Infants	Age at																						
born in	immunization	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan
NmAb	**(months)**																						
April	7								7	8	9	10	11	12	13	14	15	16					
May	6								6	7	8	9	10	11	12	13	14	15					
June	5								5	6	7	8	9	10	11	12	13	14					
July	4								4	5	6	7	8	9	10	11	12	13					
August	3								3	4	5	6	7	8	9	10	11	12					
September	2								2	3	4	5	6	7	8	9	10	11					
October	1								1	2	3	4	5	6	7	8	9	10					
November	0								0	1	2	3	4	5	6	7	8	9					
December	0									0	1	2	3	4	5	6	7	8	9				
January	0										0	1	2	3	4	5	6	7	8	9			
February	0											0	1	2	3	4	5	6	7	8	9		
March	0												0	1	2	3	4	5	6	7	8	9	
MI (seasonal)																							
November	0								0	1	2	3	4	5	6	7	8	9					
December	0									0	1	2	3	4	5	6	7	8	9				
January	0										0	1	2	3	4	5	6	7	8	9			
February	0											0	1	2	3	4	5	6	7	8	9		
March	0												0	1	2	3	4	5	6	7	8	9	
MI (year-round)																							
April	0	0	1	2	3	4	5	6	7	8	9												
May	0		0	1	2	3	4	5	6	7	8	9											
June	0			0	1	2	3	4	5	6	7	8	9										
July	0				0	1	2	3	4	5	6	7	8	9									
August	0					0	1	2	3	4	5	6	7	8	9								
September	0						0	1	2	3	4	5	6	7	8	9							
October	0							0	1	2	3	4	5	6	7	8	9						
November	0								0	1	2	3	4	5	6	7	8	9					
December	0									0	1	2	3	4	5	6	7	8	9				
January	0										0	1	2	3	4	5	6	7	8	9			
February	0											0	1	2	3	4	5	6	7	8	9		
March	0												0	1	2	3	4	5	6	7	8	9	

	Age at immunization
	Age groups (months) immunized during RSV season
	Age groups (months) immunized out off RSV season
	Age groups (months) 1 year or older of age immunized

Figure 3.1: Immunization schedules for interventions. A seasonal with catch-up strategy for NmAb, and seasonal and year-round strategies for MI were considered. Dark blue boxes represent the age at immunization, followed by subsequent months of protection during the RSV season (light blue) and out of season (light green).

2.6.4 Impact simulation

The potential impact of the implementation of an immunization program with NmAb or MI was assessed as the predicted percentage reduction in RSV-bronchiolitis and RSV-bronchiolitis hospitalizations compared to the predicted number in the absence of each preventive strategy. Also, the predicted number of the averted episodes (including 95% credible intervals) were reported. Impact was estimated in the overall study period (Nov 2010 – Jan 2016) and by RSV-season (from October to September). We simulated the evolution of RSV-bronchiolitis episodes and RSV-bronchiolitis hospitalizations from November 2010 (time t=1 in the analysis) to January 2016 (time T=270) within the

different scenarios described in Figure 1. Let $\left\{P_t^{j(m)}\right\}_{m=1}^{M}$ be the sample simulated from the posterior distribution of the model parameters without the immunization program.

For the different age groups included in the immunization schedules (either NmAb or MI), the numbers of bronchiolitis episodes, $\hat{y}_t^j$, at week t, were predicted as the mean of the M simulated values:

$$\hat{y}_t^{j(m)} = (1 - \pi_{\text{im}}) \cdot N_t^j \cdot P_t^{j(m)}$$

Where, π_{im}, was the product between the uptake and the efficacy and N_t^j was the population at time t in each age-group. The 2.5 and 97.5 percentiles of the resulting simulations were used to generate 95% prediction intervals.

For infants during unprotected months, the episodes were estimated as the product of the population and the probability of bronchiolitis episodes predicted by our model:

$$\hat{y}_t^{j(m)} = N_t^j \cdot P_t^{j(m)}$$

that is, the mean of the Binomial distribution.

The potential impact of prevention strategies against RSV was measured as the percentage reduction in potential and hospitalized RSV-bronchiolitis after implementation compared to the predicted number with no prevention strategies. Also, the numbers of the averted episodes (including the 95% predictive interval) were reported.

3. Results

3.1 Data description

A total of 64,137 all-cause bronchiolitis were registered in the cohort of children under 2 years, 36,159 episodes were considered RSV-bronchiolitis. Of the 8,142 all-cause hospitalized cases, 5,048 (62%) were lab-confirmed (RSV-bronchiolitis hospitalizations).

Figure 3.2 displays the time-series plot of the number of all-cause bronchiolitis, RSV-bronchiolitis, and RSV-bronchiolitis hospitalizations in children <2 years old. The 14.0% of RSV-bronchiolitis are hospitalizations (Table 3.3).

There was a marked seasonality, with 75% of all-cause bronchiolitis diagnosed between December and January, as well as 76% of RSV-bronchiolitis and 97% of RSV-bronchiolitis hospitalizations.

Most cases of RSV-bronchiolitis occurred in the first year of life (82%) mainly between 1 to 7 months of age. 68% of RSV bronchiolitis hospitalizations occurred in the first 3 months of life and more than 90% in the first 6 months. In infants <1 year old, hospitalizations represented 16.5% of the RSV-bronchiolitis (Table 3.3).

3.2 Impact simulation of Immunization Strategies

The model estimated that 29,872 RSV-bronchiolitis and 4,924 RSV-bronchiolitis hospitalizations occurred from November 2010 to January 2016 in children aged 0-11 months. That represents an incidence of 15,561 (95% Credibility interval (CrI): 15,215, 15,909) RSV-bronchiolitis cases per 100,000 infants-year, and 2,482 (95%CrI 2,336, 2,642) RSV-bronchiolitis hospitalizations per 100,000 infants-year.

Table 3.4 shows the impact of NmAb by coverage levels. Assuming NmAb coverages from 55% to 95%, between 10,347 (95%CrI: 10,220, 10,464) and 17,873 (95%CrI: 17,653, 18,074) RSV-bronchiolitis cases could have been averted during the study period, representing a reduction of 35% to 60%. The sensitivity analysis incorporating a 20% reduction in effectiveness showed a reduction ranging from 28% to 48%. Additionally, between 2,021 (95%CrI: 1,967, 2,082) and 3,491 (95%CrI: 3,398, 3,596) RSV-bronchiolitis hospitalizations could have been prevented, which corresponds to a reduction of 41% to 71% (see Table 3.5A). Assuming a 20% reduction of the effectiveness, the impact decreased from 31% to 53% (see annex 3.4 table 4A). Annually, between 1,016 (95%CrI: 954, 1,083) and 1,755 (95%CrI: 1,648, 1,870) RSV-bronchiolitis hospitalizations per 100,000 infants could have been averted (see Table 3.4B).

Age (months)	RSV-bronchiolitis hospitalizations N (%)	RSV-bronchiolitis N (%)
0-11 months	**4924 (97.5%)**	**29869 (82.6%)**
0	919 (18.3%)	1279 (3.5%)
1	1544 (30.7%)	2903 (8%)
2	954 (19.0%)	3129 (8.7%)
3	553 (11.0%)	3204 (8.9%)
4	329 (6.5%)	3741 (10.3%)
5	223 (4.4%)	3431 (9.5%)
6	123 (2.4%)	3281 (9.1%)
7	96 (1.9%)	2538 (7%)
8	64 (1.3%)	2085 (5.8%)
9	52 (1.0%)	1692 (4.7%)
10	50 (1.0%)	1469 (4.1%)
11	17 (0.3%)	1117 (3.1%)
12-23 months	**124 (2.5%)**	**6290 (17.4%)**
0-23 months	**5048**	**36159**

Table 3.3: Distribution of RSV-bronchiolitis hospitalizations and RSV-bronchiolitis cases by age. Cases are summarized by frequencies (N) and proportions (%).

Table 3.5 shows the impact of two scenarios (seasonal and year-round) of MI by vaccination coverage. Assuming coverages ranging from 55% to 95% in a seasonal program, between 8% and 13% of RSV-bronchiolitis and from 19% to 34% of hospitalizations would have been averted. Regarding the year-round program, between 6,492 (95%CrI: 6408, 6569) and 11,213 (95%CrI: 11,068, 11,346) RSV-bronchiolitis cases could be averted, representing a reduction of 22% to 38%. The impact decreased to a range of 16% to 27% when effectiveness was reduced by 20% (see annex 3.4 table 5A). Also, between 1473 (95%CrI: 1433, 1516) and 2544 (95%CrI: 2475, 2618) RSV-bronchiolitis hospitalizations could be prevented, which corresponds to a reduction of 30% to 52% (see Table 3.4A). A 20% effectiveness reduction showed a range of 20% to 34% (see annex 3.4 table 5A). This represents a range from 489 (95%CrI: 458,522) to 845 (95%CrI: 791, 903) RSV bronchiolitis hospitalizations in a seasonal strategy and 742 (95%CrI: 698, 790) to 1282 (95%CrI: 1206, 1365) with a year-round strategy prevented per-100000 infants-year.

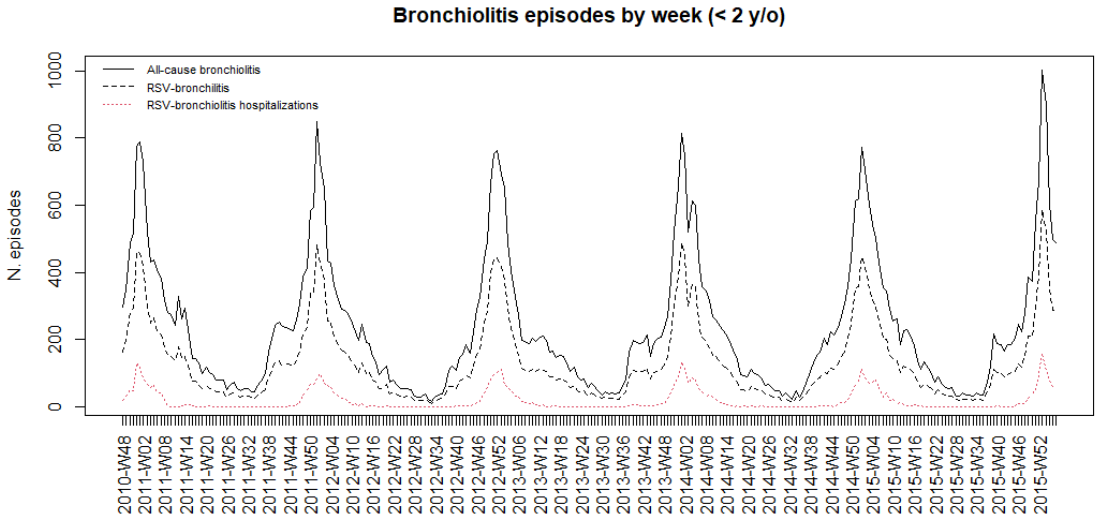


Figure 3.2: Weekly counts of bronchiolitis in children <2 y/o from November 2010 (week 48, 2010) to January 2016 (week 4, 2016). Black line: All-cause bronchiolitis. Dashed line: RSV-bronchiolitis. Red line: RSV-bronchiolitis hospitalizations.

Annex 3.3 shows the results for the sensitivity analysis of the effectiveness. Supplementary material 5 displays the weekly average counts of RSV-bronchiolitis and RSV-bronchiolitis hospitalizations predicted by the model without immunization strategies and under a catch-up and seasonal NmAb (annex 3.5a) and a year-round MI strategy (annex 3.5b). Supplementary 6 represents the impact by month of birth for the NmAb (annex 3.6a) and for the year-round MI (annex 3.6b).

4. Discussion

This paper provides the potential impact of the immunization programs with nirsevimab (Beyfortus®) or MI (Abrysvo®) on RSV-bronchiolitis and their hospitalizations, given a range of coverage using a multivariate age-structured Bayesian model. This modelling study not only used the most recent clinical trial data in the assumptions, but it was also fitted with population-based real-world data from VID (24), enhancing the validity of the results. An additional novelty would be the assessment of impact on PC, in addition to hospitalizations. Our results suggest that both strategies would effectively reduce RSV-bronchiolitis in a maximum range of 38% to 59% in RSV bronchiolitis cases and from 52% to 70% in hospitalizations depending on the strategy and program. However, the difference in endpoints, populations and timeframes between trials can all have an impact

on the efficacy values. Furthermore, the uncertainty given by the 95% credibility interval is minimal.

A) NmAb Impact estimations (Nov 10 – Jan 16)

RSV-bronchiolitis	**29872 (29533, 30198) ***				
	N (95% CrI) averted bronchiolitis; impact %				
Program/ Uptake	**55%**	**65%**	**75%**	**85%**	**95%**
Catch-up & Seasonal	10347 (10220,10464) 35%	12229 (12079,12367) 41%	14110 (13937,14269) 47%	15992 (15795,16172) 54%	17873 (17653,18074) 60%
RSV-bronchiolitis hospitalizations		**4924 (4794, 5072) ***			
Catch-up & Seasonal	2021 (1967,2082) 41%	2389 (2325,2461) 49%	2756 (2682,2839) 56%	3124 (3040,3218) 63%	3491 (3398,3596) 71%

B) NmAb Impact estimations annually from Oct -Sep **

RSV-bronchiolitis	**15561(15215, 15909) ***				
	N (95% CrI) averted bronchiolitis per 100,000 children-yr; impact %				
Program/Uptake	**55%**	**65%**	**75%**	**85%**	**95%**
Catch-up & Seasonal	5332 (5204,5460) 34%	6302 (6150,6453) 40%	7272 (7097,7446) 47%	8241 (8043,8438) 53%	9210 (8989,9431) 59%
RSV-bronchiolitis hospitalizations		**2482(2336,2642) ***			
Catch-up & Seasonal	1016 (954,1083) 41%	1201 (1127,1280) 48%	1386 (1301,1477) 56%	1570 (1474,1674) 64%	1755 (1648,1870) 70%

*Table 3.4: Impact of seasonal with catch-up NmAb program on RSV-bronchiolitis and RSV-bronchiolitis hospitalizations. A) provides the average number of bronchiolitis cases averted with the corresponding impact percentages, for the whole study period and B) by 100,000 infants per year. Different uptake MI levels (55% to 95%) were assessed. 95% CrI: 95% credible interval. * Number of predicted bronchiolitis without immunization program. ** The annual (October-September) average of number of averted bronchiolitis per 100,000 children was estimated excluding the incomplete periods 1) 2010-Nov, 2011-Sep and 2) 2015-Oct, 2016-Jan.*

A) NmAb Impact estimations (Nov 10 – Jan 16)

RSV-bronchiolitis	**29872 (29533, 30198) ***				
	N (95% CrI) averted bronchiolitis; impact %				
Program/ Uptake	**55%**	**65%**	**75%**	**85%**	**95%**
Seasonal	2243 (2205,2277) 8%	2650 (2606,2692) 9%	3058 (3007,3106) 10%	3466 (3407,3520) 12%	3874 (3808,3934) 13%
Year-round	6492 (6408,6569) 22%	7672 (7573,7763) 26%	8853 (8738,8957) 30%	10033 (9903,10152) 34%	11213 (11068,11346) 38%
RSV-bronchiolitis hospitalizations	**4924 (4794, 5072) ***				
Seasonal	956 (925,989) 19%	1130 (1093,1168) 23%	1304 (1262,1348) 26%	1478 (1430,1528) 30%	1652 (1598,1708) 34%
Year-round	1473 (1433,1516) 30%	1741 (1693,1791) 35%	2008 (1954,2067) 41%	2276 (2214,2342) 46%	2544 (2475,2618) 52%

B) NmAb Impact estimations annually from Oct -Sep **

RSV-bronchiolitis	**15561(15215,15909) ***				
	N (95% CrI) averted bronchiolitis per 100,000 children-yr; impact %				
Program/Uptake	**55%**	**65%**	**75%**	**85%**	**95%**
Seasonal	1195 (1164,1228) 8%	1413 (1375,1451) 9%	1630 (1586,1674) 10%	1847 (1798,1898) 12%	2065 (2010,2121) 13%
Year-round	3379 (3301,3457) 22%	3994 (3901,4086) 26%	4608 (4501,4714) 30%	5222 (5101,5342) 34%	5837 (5702,5971) 38%
RSV-bronchiolitis hospitalizations		**2482(2336,2642) ***			
Seasonal	489 (458,522) 20%	578 (541,617) 23%	667 (624,712) 26%	756 (707,808) 30%	845 (791,903) 34%
Year-round	742 (698,790) 30%	877 (824,934) 35%	1012 (952,1077) 41%	1147 (1078,1221) 46%	1282 (1206,1365) 52%

*Table 3.5: Impact of seasonal and year-round maternal immunization programs on RSV-bronchiolitis and RSV-bronchiolitis hospitalizations. A) provides the average number of bronchiolitis cases averted with the corresponding impact percentages, for the whole study period and B) by 100,000 infants per year. Different uptake MI levels (55% to 95%) were assessed. 95% CrI: 95% credible interval. * Number of predicted bronchiolitis without immunization program. ** The annual (October-September) average of number of averted bronchiolitis per 100,000 children was estimated excluding the incomplete periods 1) 2010-Nov, 2011-Sep and 2) 2015-Oct, 2016-Jan.*

This study shows that an immunization program with NmAb would prevent from 5,332 to 9,210 RSV-bronchiolitis and from 1,016 to 1,755 RSV-hospitalizations per 100,000 infants-year. The NmAb schedule in the model aligns with the Draft Interim Clinical Considerations of the CDC (102), ensuring that all children born from April to the end of the season receive immunization. The program suggests an intermittent administration of the mAb in hospitals, or early after discharge, and a catch up where children age < 8 months had to be immunized in a short period, coinciding with the flu vaccination in older babies and children. Hospital administration does not seem a concern, the other available mAb, palivizumab, is extensively used in hospitals before and during the RSV season. It should be noted that infant vaccination coverage in Spain exceeds 95% (109). Therefore, the NmAb program could achieve high coverage rates. In addition, preliminary data of the pragmatic clinical trial, HARMONIE, have shown higher efficacy against RSV-LRTI hospitalization (83.21%) and against very severe RSV-LRTI (75.71%) (110).

Regarding MI, the seasonal program shows a low impact on the epidemiology of bronchiolitis. It is also a challenge due to the difficulty of immunizing mothers when they are expected to deliver just before or during the RSV season. However, as shown here, a year-round immunization program would prevent between 3,379 and 5,837 RSV-bronchiolitis and 742 to 1282 RSV-hospitalizations per 100,000 children-year, and would be easier to implement than the seasonal program, though it may require further visit as it is recommended not to be administered concomitantly with the dTaP vaccine, scheduled in Spain (111).

The MATISSE trial of the MI (97,112) stopped recruitment, as recommended by the data monitoring committee, when one of the primary endpoints reached the success criterion for vaccine efficacy, which occurred before the planned number of cases were collected, and this could influence the results for other endpoints. In that trial the efficacy for MA-LRTI did not meet the success criterion (a lower boundary of the 99.5% CI greater than 20% at 90 days and 97.58% CI at later intervals) at 90 days of follow up s. Although the trial assessed all-cause RSV hospitalizations, this may include cases where infants experienced upper respiratory congestion or feeding difficulties without lung involvement. Consequently, the vaccine effectiveness could be lower in such cases. However, due to the absence of published data, the vaccine's effectiveness in preventing hospitalizations was considered based on the trial description.

Other model-based evaluations, including ordinary differential equations, deterministic dynamic models, static cohort models, decision tree models, and individual-level transmission models, also concluded that the NmAb and the MI could have great impacts on RSV disease (60,77,99,100). Additionally, Bayesian stochastic compartmental models have been also used to simulate the effect of potential immunization strategies (80,98). However, many of these evaluations were conducted prior to the availability of recent clinical trials data, using a wide range of different assumptions (target population, uptake rates, effectiveness, and duration of immunity). Consequently, the estimates obtained from these studies exhibited considerable variation. In contrast, our model shown here incorporated 1) more realistic and updated assumptions from the latest NmAb and MI results, 2) smaller-range age groups within the first year of life, allowing more precise estimates and gaining deeper insights into RSV-bronchiolitis dynamics, and 3) both population-based PC and hospital data, whereas many previous studies have focused primarily on the hospital setting. Furthermore, our modelling approach combines ideas from compartmental models and Bayesian hierarchical models in a novel and simple way explained elsewhere (98).

Nevertheless, several limitations should be noted. We assumed a constant percentage of bronchiolitis cases attributable to RSV across all calendar weeks, whereas this percentage may vary depending on the virus's circulation patterns and other factors. Also, the model did not distinguish the impact of MI on late preterm babies, whose protection might be lower, or the need for NmAb use in early preterm babies whose mothers either could not be immunized or did not have sufficient time to develop a complete immune response. Therefore, the MI program would require an NmAb catch-up for preterm infants that was not considered in the model and potentially overestimated the impact of the MI. Further, the epidemiology of bronchiolitis was characterized when palivizumab was widely used in early preterm babies. Therefore, if palivizumab is no longer used, the number of bronchiolitis cases not prevented by the MI would increase. A major limitation of this study is the limited strong data on the duration of protection of the interventions. Also, it is important to note that the conducted trials were optimized to demonstrate the efficacy of interventions in highly selected patient groups, which in some instances may limit their generalizability to a larger clinical context (23). For instance in the MI trial, a highly selected pregnant population was studied. Large trials often stop recruitment once they achieve one of their primary objectives, potentially resulting in a lack of statistical significance for some or all of the other objectives (112) in these cases we considered that the intervention was efficacious though one of the primary endpoints was not

met in the trial. A pragmatic trial (HARMONIE) has been developed with nirsevimab to provide further evidence (114). Once that trial is completed, we may need updated estimates of the impact of this preventive measure.

In conclusion, this modelling study predicts that each prophylactic strategy assessed will have an impact on the burden of bronchiolitis. Depending on the strategy and program used, we could expect an annual reduction of RSV-bronchiolitis cases by 8-59% and RSV-bronchiolitis hospitalizations by 19-70%. It is important to note that the results of the model are descriptive. However, the study used real world data and the best available evidence about each product to date. The models should be reassessed as more data of each product becomes available.

Acknowledgments

We would like to acknowledge Dr. José Antonio Lluch (Head of the Health Promotion and Prevention in Life Stages Service at the Directorate-General for Public Health, Valencia Region, Spain), Dr. José Antonio Navarro (Honorary Permanent Consultant of the Vaccines Area of the Spanish Ministry of Health), Dr. Jaime Jesús Pérez (Senior Public Health Technician of vaccines of the Murcia Region and president of the Spanish Association of Vaccinology) and Dr. David Moreno (Director of the Strategic Plan for Vaccinations in Andalusia, Spain) for their contributions designing the potential immunization programs.

The content of this chapter is under review in Journal of Infection and Public Health.

Chapter 4

A Bayesian spatiotemporal model for cluster detection: Identifying HPV suboptimal vaccine coverage.

Abstract

Human papillomavirus (HPV) is a common sexually transmitted virus responsible for several types of cancer. HPV vaccines have been included in the systematic vaccination of the Valencia Region since 2008. Despite clinical agreement on the safety and effectiveness of the vaccines, vaccination coverage remains suboptimal in many areas. In order to facilitate the implementation of targeted strategies to enhance vaccination coverage, we develop here a spatiotemporal clustering model that identifies groups of health districts that share similar behaviour. Namely, health districts are clustered twice. First, they are grouped into spatial clusters based on their underlying vaccination coverage, which is constant over time. Second, they are clustered according to their evolution of annual vaccination rates. A feature of our model is that it does not impose that geographically neighbouring areas are assigned to the same spatial cluster or the same temporal trend. This flexibility allows us to explore different spatial and temporal structures. Suboptimal HPV vaccination coverage was found in some health districts, whose coverage was around 9% lower than others.

1. Background

Human papillomavirus (HPV) is a sexually transmitted virus causative of several cancers (cervical, oropharyngeal, penile, vaginal, vulvar, and anal cancer). Among the more than 200 related types, HPV-16 and HPV-18 are the two primary high-risk types responsible for almost 70% of cervical cancers. HPV types 31, 33, 45, 52, and 58 together account for 15% of cervical cancers. HPV-6 and HPV-11 are the two primary low-risk types of causative of anogenital warts (GW) (115).

Three HPV vaccines are currently licensed: a quadrivalent vaccine (HPV 6, 11, 16, and 18) (HPV4v) available since 2006, which was followed by license of the bivalent vaccine (HPV 16 and 18) (HPV2v), and most recently the nonavalent vaccine (HPV 6, 11, 16, 18, 31, 33, 45,

52, and 58) (HPV9v). Efficacy (116–118) and effectiveness (119–122) of these vaccines preventing HPV infection, GW, precancerous lesions and cervical cancer (123,124) have been proven in different clinical trials and post-authorization studies.

In October 2008, the Valencia Region of Spain started a three-dose HPV4v vaccine programme for girls aged 14 years old. In 2011, the immunization program changed the HPV4v vaccine for the HPV2v vaccine. In 2015, the vaccination was extended to girls aged 12 years old. Since October 2022, the HPV vaccination programme included the HPV9v vaccine (for both girls and boys aged 12 years old). Although HPV vaccines are funded by the Valencian Public Health System, the uptake from 2017 to 2022 ranged between 83% and 89% (125), which is below the coverage levels achieved with other paediatric vaccines. The limited HPV vaccine uptake is probably related to two cases of vaccine-unrelated adverse events in 2008 that negatively influenced the acceptance of the vaccine, leading to an immediate decrease of 12% in the vaccination coverage (126). The study of both the spatial heterogeneity in HPV vaccination coverage at health district level and its temporal evolution would allow the identification of areas with persistently low vaccination over time. This would facilitate the implementation of targeted strategies to optimize resource allocation and enhance vaccination coverage.

Statistical models have already been applied to study the spatial variation in HPV vaccination uptake. For instance, a multilevel spatial logistic regression model was implemented to explain vaccination rates in a region of Minnesota (127). The model adjusts for individual and area-level sociodemographic characteristics and includes spatial random effects modelled as an approximated Gaussian field. A Bayesian hierarchical logistic regression model was also proposed to investigate spatial heterogeneity of HPV vaccination uptake in Switzerland (128). Spatial autocorrelation was modelled using the Besag-York-Mollié conditional autoregressive model. In (129), Bayesian hierarchical spatial regression models were used to describe the variation in HPV vaccination rates among boys and girls within ZIP codes in Virginia. Smoothed vaccination probabilities for each sex were also estimated using a conditional autoregressive (CAR) prior for spatial random effects. Correlation between sexes was then considered using either a multivariate CAR prior or the shared component model. A spatiotemporal analysis to identify priority areas in Washington State with low HPV vaccine rates over time is presented in (130). The study was performed using the CARBayesST package, which is a software package that provides a suite of models for capturing

spatiotemporal autocorrelation via random effects that are assigned CAR priors within a hierarchical Bayesian model (131).

The models discussed above are well-known models in the field of disease mapping, where the main objective is to investigate geographic disease variation across the predefined study region. By borrowing strength from nearby locations, these models provide smoothed risk surfaces and improve local estimates (132–134). However, when the objective is to identify areas exhibiting elevated (or low) risks compared with their geographical neighbours (in our scenario detection of areas with low vaccination coverage), these spatial models may fail due to oversmoothing (135). On those occasions, the use of specific methodology to address cluster detection may be advisable (136–139). In (53), the authors propose a spatial model that simultaneously allows for risk estimation and cluster identification. Unlike previous formulations, the main novelty of that model is that it does not impose any spatial correlation, so geographically separated small areas with a similar risk can be grouped in the same class. This feature allows the user to explore different risk structures.

In the space-time domain, one may aim to detect groups of contiguous areas with unusually high risk within a particular time period. Recent proposals use dummy variables to represent potential clusters and cluster detection is done by performing variable or model selection. The increase in disease risk within each cluster can be measured by including cluster-specific effects. For instance, (140) describe a model-based approach that links generalized linear models (GLM) and the spatial scan statistic (141). In particular, the approach involves fitting many different GLMs for which dummy variables that represent possible clusters are included one at a time. Cluster detection is based on selecting a number of dummy cluster variables using different criteria for model selection. An agglomerative hierarchical clustering algorithm is proposed in (142) to map out the set of potential clusters. In a second stage, for each particular configuration, a Bayesian spatiotemporal model is fitted to the data and the optimal cluster configuration is selected based on a measure of goodness-of-fit. A Lasso-based approach is proposed in (143) to detect multiple spatiotemporal clusters.

A different clustering paradigm can be found in (144), where the authors propose a Bayesian spatiotemporal mixture model for clustering areas based on their temporal trends. The candidate trend functions included in the model have fixed parametric forms or constrained shapes. This flexible methodology allows the user to test specific hypotheses about the data

being modelled while avoids identifiability problems. More recently, a Bayesian hierarchical model, which can be seen as a Gaussian Markov random field, has been proposed both to estimate spatiotemporal dependencies and to identify groups of small areas sharing similar features. The nonparametric Dirichlet process prior is employed to perform clustering of the areal-specific time series through the effects of time- and areal-specific predictors as well as CAR-modelled random effects (145).

In this chapter, we propose an alternative version of the spatial model presented in (53). This formulation can be easily implemented in statistical software for Bayesian analysis using MCMC methods. The proposed model provides estimates of the probability of vaccination in each health district in the Valencia Region, while avoiding oversmoothing. By identifying groups of areal units that share similar coverage, it enables uncovering health inequalities across the health districts. We then extend the model to the spatiotemporal domain following the guidelines in (144). The clustering of the areal-specific time series based on their trends allows the identification of groups of health districts with a similar evolution of their annual vaccination rates. This double grouping of the small areas is of great practical interest since it can be used to inform future planning efforts for geographically targeted interventions and assess the impact of vaccination policies.

2. The data

2.1 Population of interest and setting

The study included all girls born in the Valencia Region between 1995 and 2006 covered by the Public Health System (PHS). The study period covered from 01/01/2009 to 12/31/2017. During this period, target population of regional HPV vaccination programme was 14-year-old girls from October 2008 to 2014 and 12-year-old girls since 2015.

The Valencia Region is one of the 17 autonomous regions of Spain. It has a population of around 5 million inhabitants, of which 98% are covered by the Spanish PHS. The Regional Health System is divided into 241 health care districts, which are structured into 24 health departments that include at least 1 hospital, 1 specialty center and several primary care centers (241 throughout the region, corresponding to the health districts).

2.2 Data sources

We use real-world data from the Valencia health system integrated database (VID), which is a set of multiple, public, population-wide electronic databases for the Valencia Region. All the information in the VID databases can be linked at individual-level through a unique personal identification code (24). In particular, we used 1) the population information system (SIP), which provides basic information on healthcare utilization and also some sociodemographic data and 2) the vaccine information system (SIV), which stores all the information on vaccination (vaccine by type, number of doses, location and administration date, etc.). It is important to emphasize that more than 90% of public and private vaccinations are registered in the VID.

Vaccination status was obtained at individual level from the SIV. A girl was considered vaccinated against HPV if she had at least one recorded dose of HPV2v, HPV4v or HPV9v. The percentage of young women vaccinated against HPV can be estimated, for each health district and year, as the number of vaccinated girls divided by the eligible population for vaccination.

3. Methodology

In this section, we first describe an alternative formulation of the spatial clustering model proposed in (53). This new formulation can be easily implemented in statistical software for Bayesian analysis using MCMC simulation methods. We then describe a spatiotemporal extension of the model following the approach proposed in (144). This extension groups together areal units that exhibit similar underlying probabilities of vaccination and also similar temporal trends.

3.1 Spatial model

Let us assume that the study region is divided in m contiguous non-overlapping small areas and let y_i represent the observation in the i-th area, $i = 1,2, \dots, m$. Here, y_i is the number of girls vaccinated against HPV in the i-th health district ($i = 1,2, \dots, m = 241$). At the first level of the model hierarchy, we assume that the observations follow a Binomial distribution:

$$y_i \sim Bin(N_i, p_i), \tag{4.1}$$

where N_i is the (finite) target population (here, the number of girls eligible to receive the vaccine) and p_i is the probability of getting the vaccine (vaccination coverage) in area i. This probability is modeled using the logit link function as:

$$logit(p_i) = log\left(\frac{p_i}{1-p_i}\right) = log(\theta_i), \tag{4.2}$$

where parameter $\theta_i = p_i/(1-p_i) \in [0, \infty)$ is a positive quantity related to the probability of getting the vaccine. From now on we will refer to these quantities as relative probabilities. Following (53), we assume that there is a piecewise constant probability surface that describes spatial variation, and so:

$$\theta_i = z_i{}'\eta, \tag{4.3}$$

where k is the number of classes (that can be viewed as clusters), $\eta = (\eta_1, \eta_2, \dots, \eta_k)'$ is the vector of latent relative probabilities, and $z_i = (z_{i1}, z_{i2}, \dots, z_{ik})'$ is the latent indicator vector that assigns the small area i to one of the k classes; that is, z_{ij} is equal to one if the relative probability for area i corresponds to η_j and zero otherwise. In order to mitigate the label switching problem common in mixture models, we propose to order the classes. In particular, we assume:

$$\eta_1 = \Lambda_1,$$
$$\eta_j = \eta_{j-1} + \Lambda_j, \quad j = 2, \dots, k$$
$$\Lambda_j \sim Ga(1,1), \quad j = 1, \dots, k$$

so that each term is equal to the previous one plus a positive quantity.

Finally, a common prior distribution for all the indicator vectors z_i, $i = 1,2, \dots, m$, is assumed, which implies that the prior probabilities of the areas belonging to every possible class are the same. In particular:

$$z_i \sim Mult(1, \alpha),$$
$$\alpha \sim Dir(a),$$

where a is assigned a particular value. We consider here the flat Dirichlet distribution; that is, a is the $k \times 1$ vector of 1s.

Note that in disease mapping studies, y_i usually represents the count of disease. The method presented here can be applied on such occasions. In that case, p_i would represent the probability of getting the disease and θ_i the relative risk in area i.

3.2 Spatiotemporal extension

In a spatiotemporal scenario, counts that evolve over time are available for each small area. Let y_{it} be the number of vaccinated girls in small area i, $i = 1,2,\dots,m$, and time t, $t = 1,2,\dots,T$ (here $i = 1,2,\dots,m = 241$ and $t = 1,2,\dots,T = 9$). Following (144), the general model is defined as:

$$\mathrm{y_{it}} \sim \mathrm{Bin}(\mathrm{N_{it}}, \mathrm{p_{it}})$$
$$logit(p_{it}) = log(\theta_i) + \sum_{s=1}^{S} \omega_{is}\, f_s(t|\gamma_s) \qquad (4.4)$$

where N_{it} and p_{it} are defined as in (4.1) but now they refer to a particular time period t. The spatial component θ_i, which is common to all time periods, is also modelled as in (4.3). The temporal component is a clustering model that assigns each small area to one of the S temporal trends: $f_1(t|\gamma_1), f_2(t|\gamma_2), \dots, f_S(t|\gamma_S)$. The number of trends ($S$) and their particular definition are chosen by the user based on the data under study. The indicator variable ω_{is} is equal to one if the temporal trend corresponding to area i fits $f_s(t|\gamma_s)$ and zero otherwise. Since $\sum_{s=1}^{S} \omega_{is} = 1$, the following multinomial prior distribution is used for $\omega_i = (\omega_{i1}, \omega_{i2}, \dots, \omega_{iS})'$:

$$\omega_i \sim Mult(1, \beta),$$
$$\beta \sim Dir(b),$$

b being the $S \times 1$ vector of 1s.

It is important to emphasize here that the small areas are clustered twice. First, they are grouped into spatial clusters based on their underlying probability of getting the vaccine (namely, based on the constant relative probability θ_i). Second, they are clustered according to their temporal trends; that is, the evolution of annual vaccination rates. Also, allocation indicators z_i and ω_i are both modeled as independent rather than spatially autocorrelated; that is, the model does not impose that geographically neighboring areas are assigned to the same spatial cluster or the same temporal trend. Hence, this mixture model allows us to explore different spatial structures and determine the clustering of areas to trends based on the data.

3.2.1 Choice of trend functions

As recommended in (144), the trends included should be different from each other in order to get a sensible cluster identification and avoid the label switching problem. Also, in this particular case study, due to the change in the vaccination strategy that took place in 2015 (when the vaccination was extended to girls aged 12 years old), the time series for the different small areas show a drop at this year, which corresponds to time period $t = 7$. This is due to

the fact that the number of girls eligible to receive the vaccine increased, but many of these girls were reluctant to get vaccinated at such a young age. Hence, we consider the nine parametric trends described in Table 4.1.

In the previous definitions, $I(x \geq 0)$ is the indicator function that takes the value one if $x \geq 0$ and zero otherwise. Parameter γ captures the drop in the mean level of the series at time period $t = 7$. As a prior distribution for this parameter, we consider the $N(0,1000)$ distribution. The other parameters γ_s modulate the linear trends. Superscript $+$ or $-$ indicates if the linear trend is constrained to be increasing or decreasing, respectively. This is achieved by means of truncated non-informative Normal prior distributions. In particular, $\gamma_{s.}^{+} \sim N(0,1000)I(\gamma_{s.}^{+} > 0)$, while $\gamma_{s.}^{-} \sim N(0,1000)I(\gamma_{s.}^{-} < 0)$.

Stable-Stable	$f_1(t\|\gamma_1) = \gamma \cdot I(t \geq 7)$
Decreasing-Stable	$f_2(t\|\gamma_2) = \gamma_{21}^{-} \cdot t \cdot I(t \leq 6) + \gamma \cdot I(t \geq 7)$
Increasing-Stable	$f_3(t\|\gamma_3) = \gamma_{31}^{+} \cdot t \cdot I(t \leq 6) + \gamma \cdot I(t \geq 7)$
Stable-Decreasing	$f_4(t\|\gamma_4) = \gamma \cdot I(t \geq 7) + \gamma_{42}^{-} \cdot (t-7) \cdot I(t \geq 7)$
Decreasing-Decreasing	$f_5(t\|\gamma_5) = \gamma_{51}^{-} \cdot t \cdot I(t \leq 6) + \gamma \cdot I(t \geq 7) + \gamma_{52}^{-} \cdot (t-7) \cdot I(t \geq 7)$
Increasing-Decreasing	$f_6(t\|\gamma_6) = \gamma_{61}^{+} \cdot t \cdot I(t \leq 6) + \gamma \cdot I(t \geq 7) + \gamma_{62}^{-} \cdot (t-7) \cdot I(t \geq 7)$
Stable-Increasing	$f_7(t\|\gamma_7) = \gamma \cdot I(t \geq 7) + \gamma_{72}^{+} \cdot (t-7) \cdot I(t \geq 7)$
Decreasing-Increasing	$f_8(t\|\gamma_8) = \gamma_{81}^{-} \cdot t \cdot I(t \leq 6) + \gamma \cdot I(t \geq 7) + \gamma_{82}^{+} \cdot (t-7) \cdot I(t \geq 7)$
Increasing-Increasing	$f_9(t\|\gamma_9) = \gamma_{91}^{+} \cdot t \cdot I(t \leq 6) + \gamma \cdot I(t \geq 7) + \gamma_{92}^{+} \cdot (t-7) \cdot I(t \geq 7)$

Table 4.1: Trends included in the spatiotemporal model.

3.3 Model selection

Identification of the number of risk classes k (here the number of latent relative probabilities that describe spatial variation in HPV vaccination uptake) can be a challenging problem. In the context of Bayesian finite mixture models, (146) propose a criterion whose performance compares favourably with alternative criteria. The criterion, which can be easily computed using standard software for Bayesian analysis, is based on the posterior distribution of the class proportions. In particular, the mixture model is estimated with a relatively large number of latent classes. The true number of classes is then chosen as the posterior mode of the number of non-empty classes, where a class is defined as empty in an MCMC iteration of the simulation procedure if the proportion of observations assigned to that class is below a certain cut-off.

In our framework, given a number of spatial clusters k, we define $\xi_j^{(c)}$ as the proportion of small areas that are assigned to cluster j, $j = 1,2,\dots,k$, in the c-th iteration of the McMC simulation procedure. For a fixed cut-off $\varepsilon > 0$, a cluster is considered to be empty if $\xi_j^{(c)} \leq \varepsilon$. Then, the number of non-empty clusters in iteration c corresponds to $|\{j: \xi_j^{(c)} > \varepsilon\}|$, where $|\cdot|$ represents the cardinality of a set. Note that this procedure allows us to detect both clusters that are actually empty for a specific iteration and clusters that are of little relevance since they barely intervene in the data fit. Additionally, we also use the Watanabe-Akaike Information Criterion (WAIC, (147)) to assess the overall fit of each model.

3.4 Model implementation

The spatiotemporal model proposed here has been implemented with the NIMBLE system (148), a novel software for Bayesian inference based on MCMC procedures.
The 4.0.3 version of the R programming language (149) has been used for the analysis. In particular, the R packages ggplot2 (150), nimble (148), rgdal (151), rgeos (152), and spdep (153) have been employed.

4. Results

4.1 Data description

A total of 213,866 young women were followed, of whom 195,482 (91.4%) were vaccinated according to the HPV vaccination program with at least one dose. The percentage of health districts with coverage below 90% ranged from 2% to 55%, depending on the year, peaking in 2015. In 2014, 10% of the health districts reached coverage below 80%, being 22% in 2015, and 11% in 2016.

4.2 Model selection

Values of k between 2 and 10 have been tested in order to determine an optimal number of spatial clusters. Table 2 shows the WAIC obtained for each of these models. The WAIC is minimized for $k = 6$, being the optimal number of spatial clusters for our dataset. Table 2 also displays the median of the number of non-empty spatial clusters estimated for different cut-off values, namely $\varepsilon \in \{0.01,0.02,0.05\}$. When $k = 2,3,4,5$, and 6 the median matches with the number of clusters specified in the model. For larger choices of k, the number of non-empty

clusters tends to be lower than the number of clusters specified. Therefore, the choice of $k = 6$ is also suitable in terms of the number of clusters that are determined as non-empty according to the criterion in (146). Henceforth, we show the results obtained for the proposed spatiotemporal model when $k = 6$.

	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$	$k = 7$	$k = 8$	$k = 9$	$k = 10$
WAIC	13812.76	15169.03	14487.32	13430.19	12783.05	12945.27	13100.19	12971.31	12945.70
$\varepsilon = 0.01$	2	3	4	5	6	7	7	8	8
$\varepsilon = 0.02$	2	3	4	5	6	7	6	7	7
$\varepsilon = 0.05$	2	3	4	5	6	6	6	6	6

Table 4.2: WAIC values obtained for the proposed spatiotemporal model, considering a number of spatial clusters k between 2 and 10, and posterior median of the number of non-empty clusters determined for cut-off values $\varepsilon \in \{0.01, 0.02, 0.05\}$.

4.3 Spatial clustering

We group the health districts into spatial clusters according to the posterior mean of the latent indicator vectors z_i. In particular, the i-th health district is assigned to the cluster j, $j = 1,2,\ldots,6$, that presents the highest posterior mean among the $\{z_{ij}\}_{j=1}^{6}$. The estimated probabilities of getting the vaccine range from 0.90 (cluster 1) to 0.99 (cluster 6). Figure 1 contains a choropleth map of the spatial cluster assigned to each health district. All clusters included a considerable number of health districts (the number of small areas assigned to each cluster ranges between 14 and 59). A greater proportion of districts were assigned to clusters 1 and 2 in the south of the Valencia Region and also around the city of Valencia. In contrast, districts located in the north of the study region and those that are located to the south of Valencia tend to be assigned to clusters 5 and 6.

Figure 4.2 shows the posterior probabilities of each health district being assigned to each of the six spatial clusters considered. Most of the districts are assigned to a single spatial cluster, suggesting that the spatial clustering process results in sets of health districts that are well separated from each other in terms of the latent coverage rate.

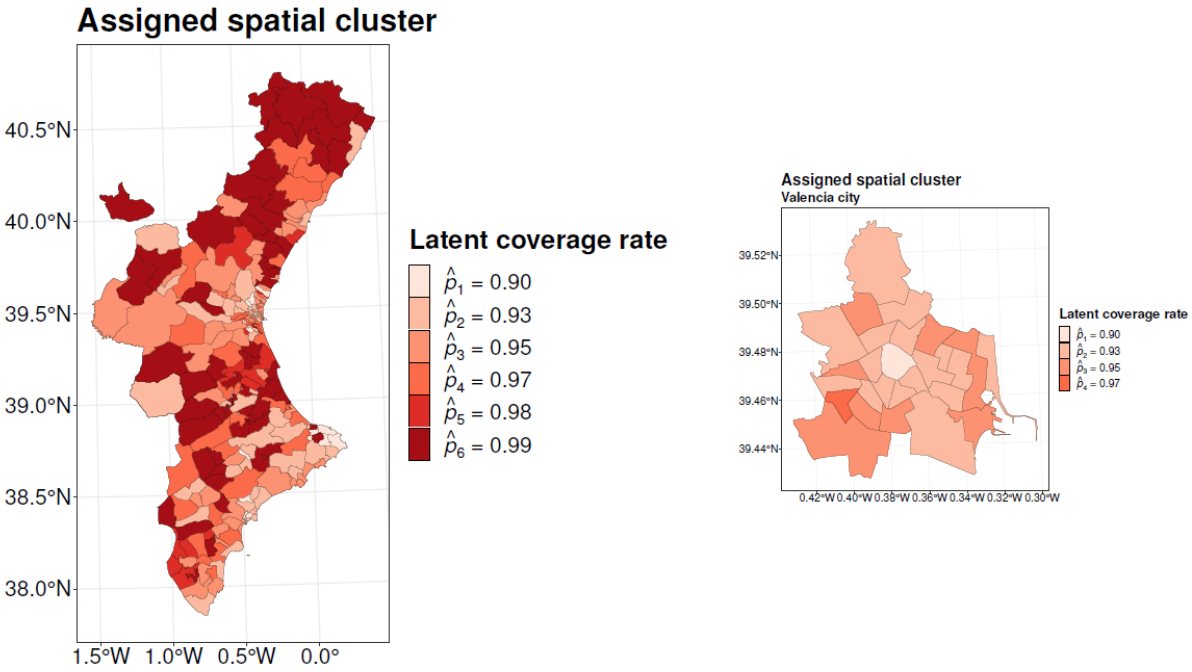


Figure 4.1: Left: Choropleth map of the spatial cluster assigned to each health district within the study region. Right: City of Valencia, expanded.

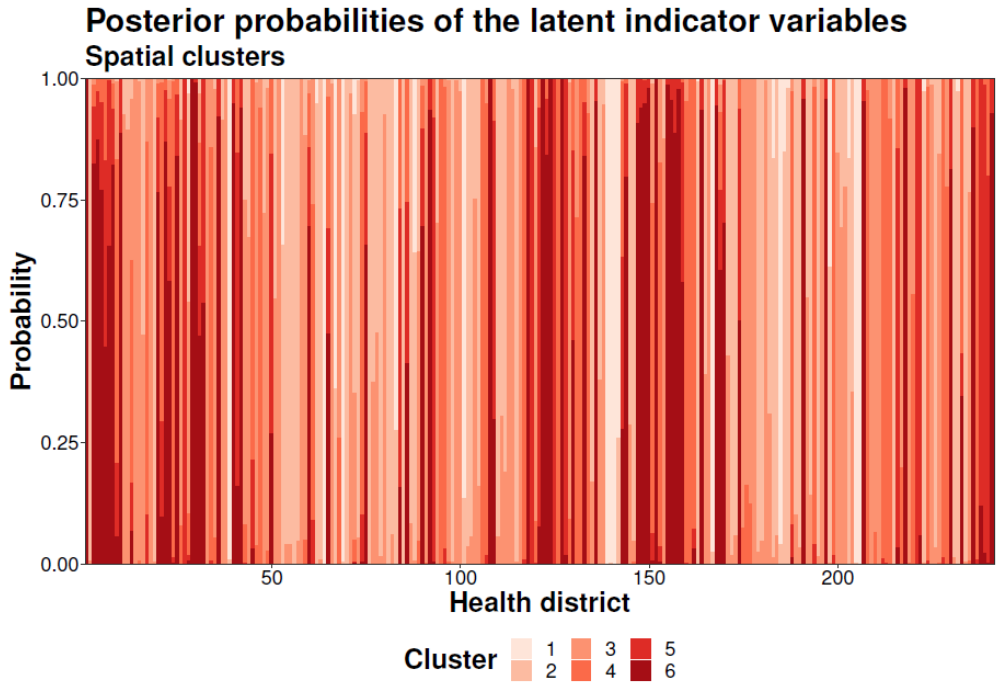


Figure 4.2: Posterior probabilities of each health district being assigned to each of the six spatial clusters considered in this analysis.

4.4 Temporal dynamics

The temporal evolution of each health district is described by the second clustering structure included in the model, which corresponds to the set of nine possible parametric trends described in Table 4.1. The nine trends include a change point at $t = 7$ that corresponds to the drop in HPV vaccination coverage that occurred in 2015 due to some changes in the vaccination strategy. This variation in the mean vaccination level is captured by the γ parameter included in the model (see Table 4.1). The posterior mean of γ is -1.28, being the 95% credible interval $(-1.33, -1.22)$.

We also assign the i-th health district to the most likely trend (temporal cluster) based on the posterior mean of the allocation variables $\{\omega_{is}\}_{s=1}^{9}$. This allows us to create the choropleth map displayed in Figure 4.3. Most of the health districts are assigned to the Increasing - Increasing and the Stable - Increasing trends (70 and 106 out of the 241 districts, respectively). No health district was assigned to the Increasing - Decreasing trend.

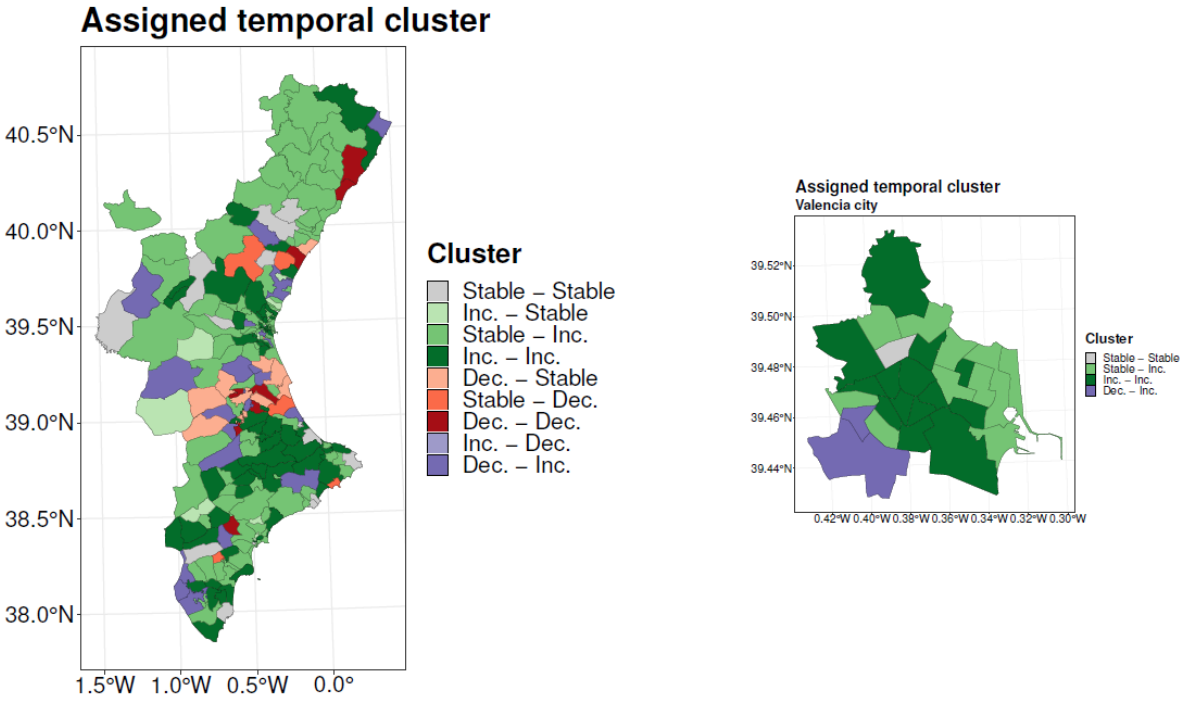


Figure 4.3: Left: Choropleth map of the time trend assigned to each health district within the study region. Right: City of Valencia, expanded.

Figure 4.4 summarizes the posterior probabilities for each health district of being assigned to each temporal cluster. Most of the districts are assigned unambiguously to a single time trend, showing robustness in the temporal clustering. In other cases, we observe a high probability of assignment for both trends Increasing - Increasing and Stable - Increasing. However, this does not seem to be a cause for concern since these trends do not differ markedly in the current study. Indeed, as can be appreciated in Figure 5, which shows the estimated annual HPV vaccination rates for the study period, these two-time trends have evolved similarly. In fact, the increasing trends identified for some health districts before 2015 are rather subtle, while the same kind of trends estimated from year 2015 are quite marked.

Finally, Table 4.3 shows the estimates of all slope parameters involved in the predefined trends.

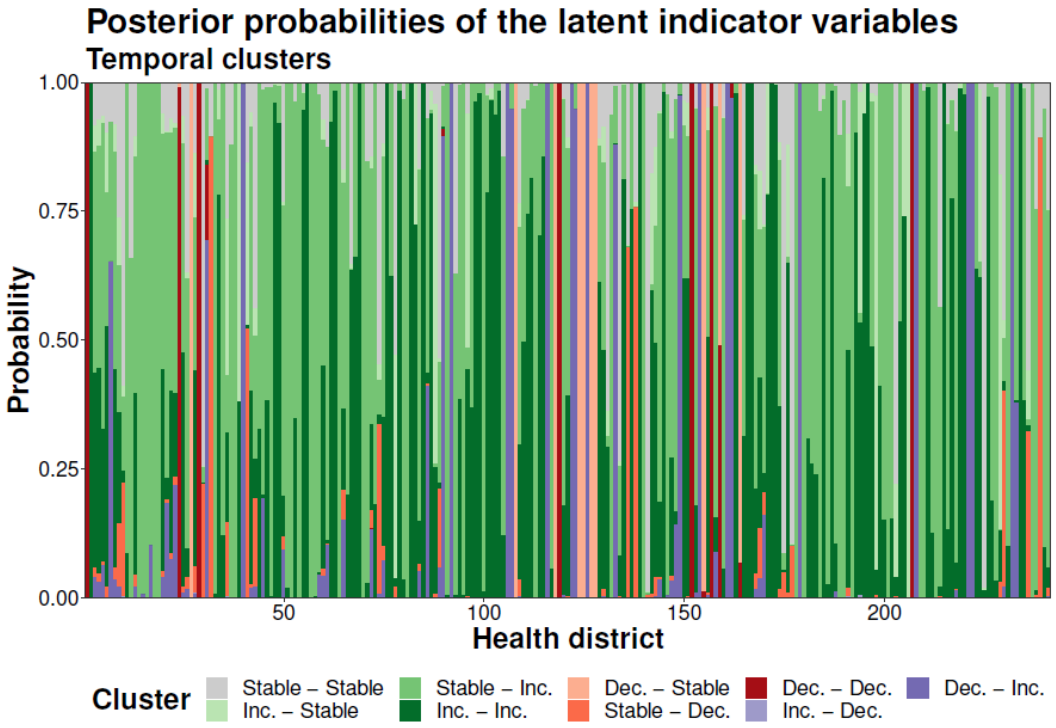


Figure 4.4: Posterior probabilities of each health district being assigned to each of the nine time trends considered.

Trend	Parameter	Mean	Lo	Up	Parameter	Mean	Lo	Up
Decreasing- Stable	γ_{21}^{-}	-1.03	-1.06	-1.00	-	-	-	-
Increasing - Stable	γ_{31}^{+}	0.32	0.17	0.50	-	-	-	-
Stable - Decreasing	-	-	-	-	γ_{42}^{-}	-0.69	-5.56	-0.21
Decreasing- Decreasing	γ_{51}^{-}	-0.75	-0.78	-0.72	γ_{52}^{-}	-0.05	-0.15	-0.00
Increasing - Decreasing	γ_{61}^{+}	25.37	1.12	71.98	γ_{62}^{-}	-25.08	-71.48	-0.65
Stable - Increasing	-	-	-	-	γ_{72}^{+}	0.64	0.58	0.70
Decreasing - Increasing	γ_{81}^{-}	-0.36	-0.38	-0.34	γ_{82}^{+}	0.37	0.26	0.49
Increasing - Increasing	γ_{91}^{+}	0.21	0.18	0.24	γ_{92}^{+}	0.65	0.57	0.73

Table 4.3: Posterior mean and 95% credible interval for the parameters involved in the parametric time trends considered in the model.

5. Discussion

We have presented a novel spatiotemporal model to identify clusters of health districts with HPV vaccination coverage lower than expected in a vaccination program. The study was based on high-quality population-based data of young women from the Valencia Region with a long-term follow-up period of nine years.

Previous studies have been developed to examine spatial and spatiotemporal variation in HPV vaccination coverage (127–130), but they do not address cluster detection. Considering the limited variability in vaccination practices incorporated into the public health schedule, the use of spatially correlated random effects to smooth HPV vaccination rates may mask subtle differences between neighbouring areas. Therefore, the use of a clustering model avoiding oversmoothing may be advisable.

The spatial model proposed here represents an improved version of the model presented in (53). The alternative formulation simplifies its implementation in statistical software for Bayesian analysis. In this way, we enhance its reproducibility, and it can be used by a greater number of practitioners. Following the methodology described in (144), the model has been extended to the spatiotemporal domain to describe temporal trends. The incorporation of temporality in the spatial model allowed us both to characterize the temporal pattern of annual vaccination rates and to detect departures from stable behaviours. The spatiotemporal clustering model identifies both spatial and temporal clusters. First, health districts are grouped based on their underlying vaccination probability, which is constant over time. Next, health districts are clustered according to their evolution of annual vaccination rates. This double grouping of the small areas is of great practical interest since it can be used to inform future planning efforts for geographically targeted interventions and assess the impact of vaccination

policies. This model could also be applied to study the uptake of other vaccines or the risk of diseases.

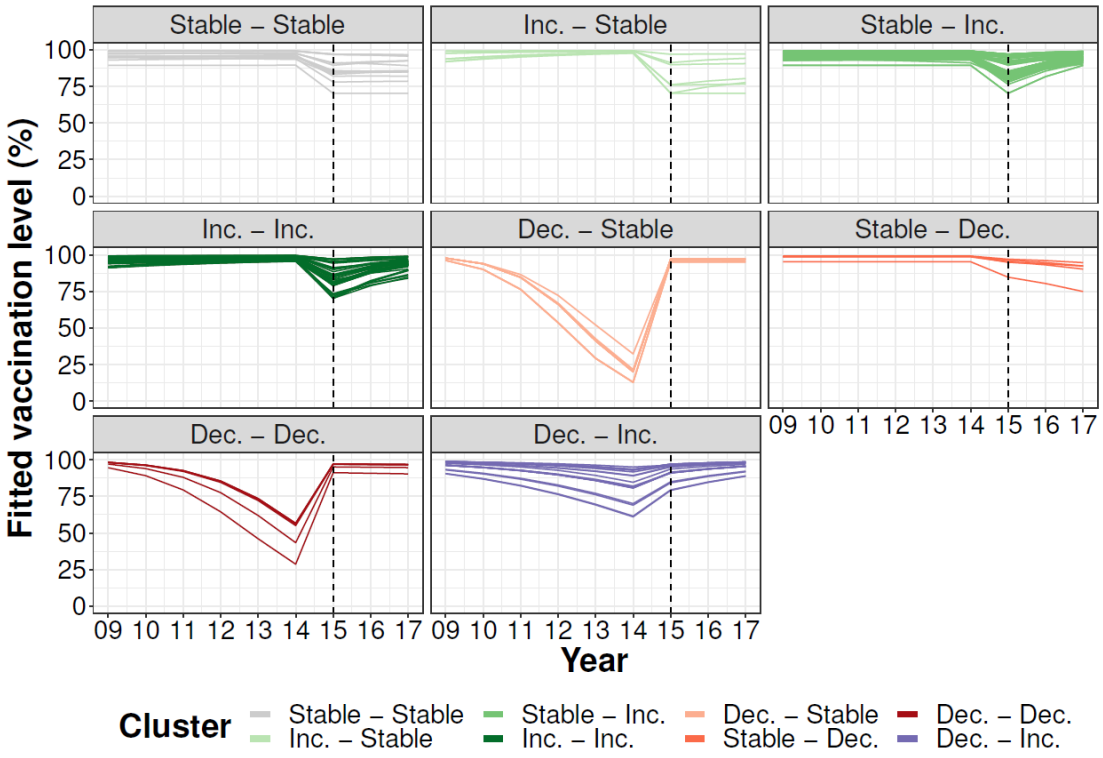


Figure 4.5: Estimated HPV vaccination coverage ($\hat{p}_{it}$) for the 241 health districts during the period 2009-2017. The estimates are grouped according to the time trend assigned to each health district. The dashed line is located at year $t = 2015$, where the change point took place.

Even though HPV vaccine is included in the Spanish vaccination calendar, and so it is free, our study provides evidence of the presence of health districts with suboptimal HPV vaccination. In fact, the results obtained reveal the presence of six spatial clusters, with coverage ranging between 90% and 99%. Most health districts were assigned to clusters 1 and 2, which correspond to the lowest levels of HPV vaccination. These areas are concentrated in the south of the Valencia Region and around the city of Valencia. On the contrary, the provinces of Alicante and Castellón have more areas assigned to clusters with higher vaccination coverage. A plausible explanation could be the negative messages in the media in the presence of two cases of adverse events consecutive to vaccine inoculation in the city of Valencia. Also, paediatrician recommendation strongly motivates parents to vaccinate their adolescent children against HPV. It is interesting to note that one of the wealthiest areas of the city of Valencia (the center of Valencia) is assigned to the lowest probability cluster. Regarding

temporal clusters, health districts were grouped into eight temporal trends. All trends include a change point in 2015 to capture the drop in HPV vaccination that occurred after including 12-year-old girls in the vaccination schedule. However, after this drop, most areas experienced a recovery. In fact, most of the health districts are assigned to either the Increasing - Increasing or Stable - Increasing categories.

This study focuses exclusively on young women. However, it is worth noting that, in the Valencia Region, systematic vaccination for boys was introduced in 2022. One of our future objectives involves the implementation of a multivariate model to assess and compare HPV vaccination rates in both sexes.

A recognized limitation of our current study is the lack of sociodemographic data, such as information on vulnerability, ethnicity, rural settings, etc. It would be interesting to extend the model to take into account covariate information, so that we can explore its relationship with HPV vaccination rates.

6. Conclusions

We provide a novel spatiotemporal model that is 1) easy to implement and 2) capable of classifying geographical areas based on the underlying spatial pattern as well as their temporal trends. Suboptimal HPV vaccination coverage was found in some health districts. Additional information on potential factors influencing vaccination uptake is needed to better understand health inequalities across the health districts.

The content of this chapter is under review in Stochastic Environmental Research and Risk Assessment (SERRA) journal.

Chapter 5

Conclusions and future research

In this thesis, we have focused on providing guidance for the prevention of two infectious diseases of significant interest both worldwide and within the Valencia Region: RSV, which is a leading cause of paediatric hospitalizations, and HPV, responsible for various types of cancers. We have contributed with enhanced statistical models to address important research questions, such as estimating the potential impact of introducing a new immunisation intervention or identifying areas with suboptimal vaccination. This information is a valuable tool to guide public health policy makers.

We have first investigated the potential impact that upcoming immunization strategies might have on the burden of RSV. Given that RSV is the primary cause of bronchiolitis, our case studies have rather focused on bronchiolitis episodes. The aim of our first research paper (which correspond to Chapter two) was to find a robust model to accurately describe the dynamics of bronchiolitis in children under two years of age, using high-quality data from the Valencia health system integrated database, in a scenario without preventive strategies. We proposed a stochastic multivariate age-structured model that, innovatively, combined ideas from compartmental models with Bayesian hierarchical Poisson models for count data. Discrete time series of bronchiolitis counts were divided into four age groups, and new infections were described by considering the interaction among these age groups. Unlike standard formulations, the main innovation of our model is that the mean of the distribution at each time point depends only on an observation-driven component, where the autoregressive parameter is allowed to vary stochastically over time. Also, we introduced a mixing parameter to account for heterogeneous mixing of individuals in the population. Our approach does not require information on susceptibles, and it provides a simplified framework to describe disease counts, avoiding the modelling of complex transitions between different compartments. We found that previous counts of bronchiolitis have a greater impact on the youngest children and that seasonality strongly influences transmission dynamics. However, given that the seasonal pattern varies slightly from year to year, incorporating random effects to account for stochasticity in transmission was crucial for providing a more accurate description of the data. Furthermore, a simulation extension of this model was performed to assess the potential impact of an immunization scenario consisting of vaccinating a proportion of newborns.

With the recent approval of the monoclonal antibody nirsevimab (Beyfortus®) and the bivalent prefusion RSV vaccine (Abrysvo®) for maternal immunization, we carried out a more precise study to support public health decisions on incorporating these immunizations into schedules. This study aimed to provide more precise and realistic impact estimations of each product on RSV bronchiolitis episodes in infants under 12 months. We considered thirteen age-groups and incorporated the new findings regarding efficacy and protection duration in clinical trials. As in the previous case study, we modelled first the age-structured time series of RSV bronchiolitis (treated in primary care and laboratory confirmed RSV hospitalizations) by a multivariate age-structured Bayesian model. Because we have smaller populations, the model assumed that the distribution governing the counts was the Binomial distribution. An extension of the model was then implemented to simulate the potential impact of each immunization strategy under different assumptions of effectiveness, duration of protection and coverage. It is important to note that this is a descriptive study. Differences in endpoints, studied populations, and timeframes between clinical trials pose challenges to directly comparing the two interventions. However, our study provides valuable insights into their potential impact.

While previous studies have shown that the new monoclonal antibody and maternal immunization could effectively reduce hospitalizations, there is limited research on their impact in ambulatory care settings. Our study offers a realistic simulation scenario that includes primary care and hospital bronchiolitis data and the last findings regarding efficacies and duration of protection. We analysed long-term data from November 2010 to the end of January 2016, including 290,484 infants, with 36,159 episodes of RSV bronchiolitis and 5,048 RSV hospitalizations. The results suggest that both strategies would effectively reduce RSV bronchiolitis in a maximum range of 5,837 to 9,210 RSV bronchiolitis cases per 100,000 infants-year and from 1,282 to 1,755 hospitalizations depending on the strategy and program.

We have also investigated suboptimal HPV vaccination coverage in the Valencia Region using a novel spatiotemporal model. Even though HPV vaccines are part of systematic vaccination since 2008, coverage remains low in many areas. We have developed a novel spatiotemporal clustering model that identifies groups of health districts that share similar behaviour. Namely, health districts were clustered twice. First, they were grouped into spatial clusters based on their underlying vaccination coverage, which was constant over time. Second, they were clustered according to their evolution of annual vaccination rates. Unlike

previous studies that explored spatial and spatiotemporal variations in HPV vaccination coverage without focusing on cluster detection, this model offers a refined approach that avoids oversmoothing issues inherent in spatial analyses, thus revealing subtle differences between neighbouring areas. Suboptimal HPV vaccination coverage was found in some health districts, whose coverage was around 9% lower than others. We highlight the need for further research and model refinement incorporating sociodemographic and socioeconomic covariates to better understand vaccination disparities.

The modelling approaches developed in this thesis have the potential to be adapted and applied to other infectious diseases, and in the case of the spatiotemporal clustering model, to other vaccination programs to assess their spatiotemporal inequalities.

Future research lines:

- RSV infections can pose significant risks for older adults, which has led to the recent approval of a vaccine for individuals over 60 years old. We plan to extend our multivariate age-structured stochastic model to include subjects over 60 years. We will also develop an extension of the model to estimate the potential impact of introducing the vaccine in older adults on RSV burden. Additionally, with the recent inclusion of nirsevimab in the schedules of the Valencia Region a few months ago, we will update our model to incorporate its effects on RSV transmission dynamics.
- Considering that socio-demographic and economic covariates could explain areas of suboptimal HPV vaccination, it would be interesting to extend our spatiotemporal model for cluster detection including these covariates in the model.
- The third case study focuses exclusively on young women. However, the Valencia Region included the systematic vaccination for boys in 2022. One of our future research projects aims to implement a multivariate model to assess and compare HPV vaccination rates in both sexes within a spatiotemporal framework.
- The immunization program with nirsevimab in the Valencia Region targets two infant groups: those born during the RSV season who are immunized at birth (seasonal), and catch-up infants aged 0 to 6 months born outside of the season. We aim to implement the spatiotemporal model to analyse nirsevimab coverage. This analysis will help to identify areas with suboptimal immunization, potentially influenced by socioeconomic and demographic factors. Additionally, we will compare the immunization rates between seasonal and catch-up infants.

Bibliography

1. White PJ. Mathematical Models in Infectious Disease Epidemiology. Infect Dis. 2017;49-53.e1.

2. Gosztonyi K. How history of mathematics can help to face a crisis situation: the case of the polemic between Bernoulli and d'Alembert about the smallpox epidemic. Educ Stud Math. 2021 Oct 1;108(1):105–22.

3. Bernoulli, Daniel - Encyclopedia of Mathematics [Internet]. [cited 2024 Jan 14]. Available from: https://encyclopediaofmath.org/wiki/Bernoulli,_Daniel

4. History of Smallpox | Smallpox | CDC [Internet]. 2021 [cited 2024 Jan 14]. Available from: https://www.cdc.gov/smallpox/history/history.html

5. Quezada A. Los orígenes de la vacuna. Rev Médica Clínica Las Condes. 2020 May 1;31(3):367–73.

6. Martinez EZ, Achcar JA. Trends in epidemiology in the 21st century: time to adopt Bayesian methods. Cad Saúde Pública. 2014 Apr;30:703–14.

7. Hamson E, Forbes C, Wittkopf P, Pandey A, Mendes D, Kowalik J, et al. Impact of pandemics and disruptions to vaccination on infectious diseases epidemiology past and present. Hum Vaccines Immunother. 2023 Aug;19(2):2219577.

8. Yadav SK, Akhter Y. Statistical Modeling for the Prediction of Infectious Disease Dissemination With Special Reference to COVID-19 Spread. Front Public Health [Internet]. 2021 [cited 2023 Dec 30];9. Available from: https://www.frontiersin.org/articles/10.3389/fpubh.2021.645405

9. Introduction to Infectious Diseases [Internet]. [cited 2023 Dec 26]. Available from: https://www.bcm.edu/departments/molecular-virology-and-microbiology/emerging-infections-and-biodefense/introduction-to-infectious-diseases

10. Joint Document of the Spanish Society of Infectious Diseases and Clinical Microbiology (SEIMC) and the Spanish Society of Preventive Medicine, Public Health and Hygiene (SEMPSPH) for Combating Antimicrobial Resistance. Enfermedades Infecc Microbiol Clínica. 2017 Mar 1;35(3):139–40.

11. Vaccines and immunization [Internet]. [cited 2023 Dec 26]. Available from: https://www.who.int/health-topics/vaccines-and-immunization

12. Smith PG. Concepts of herd protection and immunity. Procedia Vaccinol. 2010 Jan 1;2(2):134–9.

13. Haber M. Estimation of the direct and indirect effects of vaccination. Stat Med. 1999;18(16):2101–9.

14. Wicker S, Maltezou HC. Vaccine-preventable diseases in Europe: where do we stand? Expert Rev Vaccines. 2014 Aug 1;13(8):979–87.

15. Vaccine Glossary of Terms | CDC [Internet]. 2023 [cited 2023 Dec 30]. Available from: https://www.cdc.gov/vaccines/terms/glossary.html

16. Types of Immunity to a Disease | CDC [Internet]. 2022 [cited 2023 Dec 26]. Available from: https://www.cdc.gov/vaccines/vac-gen/immunity-types.htm

17. Monoclonal antibodies (mAbs) for infectious diseases [Internet]. [cited 2023 Dec 26]. Available from: https://www.who.int/teams/immunization-vaccines-and-biologicals/product-and-delivery-research/monoclonal-antibodies-(mabs)-for-infectious-diseases

18. Delany I, Rappuoli R, De Gregorio E. Vaccines for the 21st century. EMBO Mol Med. 2014 Jun;6(6):708–20.

19. López-Lacort M, Orrico-Sánchez A, Martínez-Beneito MÁ, Muñoz-Quiles C, Díez-Domingo J. Spatio-temporal impact of self-financed rotavirus vaccination on rotavirus and acute gastroenteritis hospitalisations in the Valencia region, Spain. BMC Infect Dis. 2020 Sep 7;20:656.

20. Infectious Disease Epidemiology – Public Health [Internet]. [cited 2023 Dec 30]. Available from: https://publichealth.cornell.edu/curriculum/concentrations/infectious-disease-epidemiology/

21. Scherer A, McLean A. Mathematical models of vaccination. Br Med Bull. 2002;62:187–99.

22. Overton CE, Stage HB, Ahmad S, Curran-Sebastian J, Dark P, Das R, et al. Using statistics and mathematical modelling to understand infectious disease outbreaks: COVID-19 as an example. Infect Dis Model. 2020 Jan 1;5:409–41.

23. Commissioner O of the. FDA. FDA; 2023 [cited 2024 Jan 20]. Real-World Evidence. Available from: https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence

24. García-Sempere A, Orrico-Sánchez A, Muñoz-Quiles C, Hurtado I, Peiró S, Sanfélix-Gimeno G, et al. Data Resource Profile: The Valencia Health System Integrated Database (VID). Int J Epidemiol. 2020 Jun 1;49(3):740–741e.

25. Kretzschmar M. Disease modeling for public health: added value, challenges, and institutional constraints. J Public Health Policy. 2020;41(1):39–51.

26. Adiga A, Dubhashi D, Lewis B, Marathe M, Venkatramanan S, Vullikanti A. Mathematical Models for COVID-19 Pandemic: A Comparative Analysis. J Indian Inst Sci. 2020;100(4):793–807.

27. Hong K, Kim J, Yum S, Gómez Gómez RE, Chun BC. 1442. Spatiotemporal Clusters of Varicella and the Regional Risks through Bayesian Approach: A National Five-year Cohort Analysis. Open Forum Infect Dis. 2021 Nov 1;8(Supplement_1):S802.

28. Orrico-Sanchez A, López-Lacort M, Pérez-Vilar S, Díez-Domingo J. Long-term impact of self-financed rotavirus vaccines on rotavirus-associated hospitalizations and costs in the Valencia Region, Spain. BMC Infect Dis. 2017 Dec;17(1):267.

29. Mordecai EA, Cohen JM, Evans MV, Gudapati P, Johnson LR, Lippi CA, et al. Detecting the impact of temperature on transmission of Zika, dengue, and chikungunya using mechanistic models. PLoS Negl Trop Dis. 2017 Apr 27;11(4):e0005568.

30. Shmueli G. To Explain or to Predict? Stat Sci [Internet]. 2010 Aug 1 [cited 2024 Jan 14];25(3). Available from: https://projecteuclid.org/journals/statistical-science/volume-25/issue-3/To-Explain-or-to-Predict/10.1214/10-STS330.full

31. Ashby D. Bayesian statistics in medicine: a 25 year review. Stat Med. 2006;25(21):3589–631.

32. Greenland S. Bayesian perspectives for epidemiological research: I. Foundations and basic methods. Int J Epidemiol. 2006 Jun 1;35(3):765–75.

33. Tobias JL. Primer on the Use of Bayesian Methods in Health Economics. In: Culyer AJ, editor. Encyclopedia of Health Economics [Internet]. San Diego: Elsevier; 2014 [cited 2024 Jan 28]. p. 146–54. Available from: https://www.sciencedirect.com/science/article/pii/B9780123756787007082

34. Yildirim I. Bayesian Inference: Metropolis-Hastings Sampling.

35. Štrumbelj E, Bouchard-Côté A, Corander J, Gelman A, Rue H, Murray L, et al. Past, Present, and Future of Software for Bayesian Inference.

36. Sofonea MT, Cauchemez S, Boëlle PY. Epidemic models: why and how to use them. Anaesth Crit Care Pain Med. 2022 Apr;41(2):101048.

37. Kretzschmar M, Wallinga J. Mathematical Models in Infectious Disease Epidemiology. In: Krämer A, Kretzschmar M, Krickeberg K, editors. Modern Infectious Disease Epidemiology: Concepts, Methods, Mathematical Models, and Public Health [Internet]. New York, NY: Springer; 2010 [cited 2023 Dec 27]. p. 209–21. (Statistics for Biology and Health). Available from: https://doi.org/10.1007/978-0-387-93835-6_12

38. Brauer F. Mathematical epidemiology: Past, present, and future. Infect Dis Model. 2017 May 1;2(2):113–27.

39. Brauer F. Compartmental Models in Epidemiology. Math Epidemiol. 2008;1945:19–79.

40. Seydou M, Tessa OM. A Stochastic SVIR Model for Measles. Appl Math. 2021 Mar 11;12(3):209–23.

41. Nepomuceno EG, Takahashi RHC, Aguirre LA. 2016. Individual-based model (IBM): an alternative framework for epidemiological compartment models. Biometric Brazilian J. 34, 133–162.

42. Ortega NRS, Santos FS, Zanetta DMT, Massad E. A Fuzzy Reed–Frost Model for Epidemic Spreading. Bull Math Biol. 2008 Oct 1;70(7):1925–36.

43. Allen LJS, Burgin AM. Comparison of deterministic and stochastic SIS and SIR models in discrete time. Math Biosci. 2000 Jan 1;163(1):1–33.

44. Britton T. Basic stochastic transmission models and their inference [Internet]. arXiv; 2018 [cited 2024 Mar 3]. Available from: http://arxiv.org/abs/1801.09594

45. Choisy M, Guégan JF, Rohani P. Mathematical modeling of infectious disease dynamics. In 2007. p. 379–404.

46. Bender R. Introduction to the use of regression models in epidemiology. Methods Mol Biol Clifton NJ. 2009;471:179–95.

47. Generalized linear model. In: Wikipedia [Internet]. 2023 [cited 2024 Jan 30]. Available from: https://en.wikipedia.org/w/index.php?title=Generalized_linear_model&oldid=1189494117

48. Held L, Hofmann M, Höhle M, Schmid V. A Two-Component Model for Counts of Infectious Diseases. Biostatistics. 2006;7(3):422–37.

49. Held L, Höhle M, Hofmann M. A statistical framework for the analysis of multivariate infectious disease surveillance data. Stat Modllng. 2005 Jan 1;5.

50. Corberán-Vallet A, Santonja FJ. A Bayesian SIRS model for the analysis of respiratory syncytial virus in the region of Valencia, Spain. Biom J Biom Z. 2014 Sep;56(5):808–18.

51. Bracher J, Held L. Endemic-epidemic models with discrete-time serial interval distributions for infectious disease prediction. Int J Forecast. 2022 Jul 1;38(3):1221–33.

52. Samat NA, Mey LW. Malaria Disease Mapping in Malaysia based on Besag-York-Mollie (BYM) Model. J Phys Conf Ser. 2017 Sep;890(1):012167.

53. Flórez KC, Corberán-Vallet A, Iftimi A, Bermúdez JD. A Bayesian unified framework for risk estimation and cluster identification in small area health data analysis. PLOS ONE. 2020 May 7;15(5):e0231935.

54. Korsten K, Bont L. Seasonal immunisation against respiratory syncytial virus disease. Lancet Public Health. 2017 Aug 1;2(8):e344–5.

55. Ralston SL, Lieberthal AS, Meissner HC, Alverson BK, Baley JE, Gadomski AM, et al. Clinical practice guideline: the diagnosis, management, and prevention of bronchiolitis. Pediatrics. 2014 Nov;134(5):e1474-1502.

56. Díez Domingo J, Ridao López M, Ubeda Sansano I, Ballester Sanz A. [Incidence and cost of hospitalizations for bronchiolitis and respiratory syncytial virus infections in the autonomous community of Valencia in Spain (2001 and 2002)]. An Pediatr Barc Spain 2003. 2006 Oct;65(4):325–30.

57. Muñoz-Quiles C, López-Lacort M, Úbeda-Sansano I, Alemán-Sánchez S, Pérez-Vilar S, Puig-Barberà J, et al. Population-based Analysis of Bronchiolitis Epidemiology in Valencia, Spain. Pediatr Infect Dis J. 2016 Mar;35(3):275–80.

58. Resch B. Respiratory Syncytial Virus Infection in High-risk Infants – an Update on Palivizumab Prophylaxis. Open Microbiol J. 2014 Jul 11;8:71–7.

59. Shook BC, Lin K. Recent Advances in Developing Antiviral Therapies for Respiratory Syncytial Virus. Top Curr Chem Cham. 2017 Apr;375(2):40.

60. Rainisch G, Adhikari B, Meltzer MI, Langley G. Estimating the impact of multiple immunization products on medically-attended respiratory syncytial virus (RSV) infections in infants. Vaccine. 2020 Jan 10;38(2):251–7.

61. Domachowske JB, Anderson EJ, Goldstein M. The Future of Respiratory Syncytial Virus Disease Prevention and Treatment. Infect Dis Ther. 2021 Mar;10(Suppl 1):47–60.

62. Madhi SA, Polack FP, Piedra PA, Munoz FM, Trenholme AA, Simões EAF, et al. Respiratory Syncytial Virus Vaccination during Pregnancy and Effects in Infants. N Engl J Med. 2020 Jul 30;383(5):426–39.

63. Griffin MP, Yuan Y, Takas T, Domachowske JB, Madhi SA, Manzoni P, et al. Single-Dose Nirsevimab for Prevention of RSV in Preterm Infants. N Engl J Med. 2020 Jul 30;383(5):415–25.

64. Lanari M, Giovannini M, Giuffré L, Marini A, Rondini G, Rossi GA, et al. Prevalence of respiratory syncytial virus infection in Italian infants hospitalized for acute lower respiratory tract infections, and association between respiratory syncytial virus infection risk factors and disease severity. Pediatr Pulmonol. 2002 Jun;33(6):458–65.

65. Koehoorn M, Karr CJ, Demers PA, Lencar C, Tamburic L, Brauer M. Descriptive epidemiological features of bronchiolitis in a population-based cohort. Pediatrics. 2008 Dec;122(6):1196–203.

66. García CG, Bhore R, Soriano-Fallas A, Trost M, Chason R, Ramilo O, et al. Risk factors in children hospitalized with RSV bronchiolitis versus non-RSV bronchiolitis. Pediatrics. 2010 Dec;126(6):e1453-1460.

67. Frequency, duration and predictors of bronchiolitis episodes of care among infants ≥32 weeks gestation in a large integrated healthcare system: a retrospective cohort study | BMC Health Services Research | Full Text [Internet]. [cited 2024 Feb 21]. Available from: https://bmchealthservres.biomedcentral.com/articles/10.1186/1472-6963-12-144

68. Murray J, Bottle A, Sharland M, Modi N, Aylin P, Majeed A, et al. Risk factors for hospital admission with RSV bronchiolitis in England: a population-based birth cohort study. PloS One. 2014;9(2):e89186.

69. Weber A, Weber M, Milligan P. Modeling epidemics caused by respiratory syncytial virus (RSV). Math Biosci. 2001 Aug;172(2):95–113.

70. Acedo L, Díez-Domingo J, Moraño JA, Villanueva RJ. Mathematical modelling of respiratory syncytial virus (RSV): vaccination strategies and budget applications. Epidemiol Infect. 2010 Jun;138(6):853–60.

71. Modeling the variations in pediatric respiratory syncytial virus seasonal epidemics | BMC Infectious Diseases | Full Text [Internet]. [cited 2024 Feb 21]. Available from: https://bmcinfectdis.biomedcentral.com/articles/10.1186/1471-2334-11-105

72. Modelling the Seasonal Epidemics of Respiratory Syncytial Virus in Young Children | PLOS ONE [Internet]. [cited 2024 Feb 21]. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0100422

73. Isham V. Stochastic models for epidemics. In: Celebrating Statistics: Papers in Honour of Sir David Cox on the occasion of his 80th birthday. Oxford University Press; 2005. p. 27–53.

74. Allen L. An Introduction to Stochastic Epidemic Models. In: Mathematical Epidemiology. Springer; 2008. p. 81–130.

75. Greenwood PE, Gordillo LF. Stochastic Epidemic Modeling. In: Chowell G, Hyman JM, Bettencourt LMA, Castillo-Chavez C, editors. Mathematical and Statistical Estimation Approaches in Epidemiology [Internet]. Dordrecht: Springer Netherlands; 2009 [cited 2024 Feb 21]. p. 31–52. Available from: https://doi.org/10.1007/978-90-481-2313-1_2

76. Treskova M, Pozo-Martin F, Scholz S, Schönfeld V, Wichmann O, Harder T. Assessment of the Effects of Active Immunisation against Respiratory Syncytial Virus (RSV) using Decision-Analytic Models: A Systematic Review with a Focus on Vaccination Strategies, Modelling Methods and Input Data. PharmacoEconomics. 2021 Mar;39(3):287–315.

77. Mezei A, Cohen J, Renwick MJ, Atwell J, Portnoy A. Mathematical modelling of respiratory syncytial virus (RSV) in low- and middle-income countries: A systematic review. Epidemics. 2021 Jun 1;35:100444.

78. Hogan AB, Anderssen RS, Davis S, Moore HC, Lim FJ, Fathima P, et al. Time series analysis of RSV and bronchiolitis seasonality in temperate and tropical Western Australia. Epidemics. 2016 Sep 1;16:49–55.

79. Hodgson D, Pebody R, Panovska-Griffiths J, Baguelin M, Atkins KE. Evaluating the next generation of RSV intervention strategies: a mathematical modelling study and cost-effectiveness analysis. BMC Med. 2020 Nov 18;18(1):348.

80. Jornet-Sanz M, Corberan-Vallet A, Santonja FJ, Villanueva RJ. A Bayesian stochastic SIRS model with a vaccination strategy for the analysis of respiratory syncytial virus.

81. Held L, Höhle M, Hofmann M. A statistical Framework for the Analysis of Multivariate Infectious Disease Surveillance Counts. Stat Model. 2005;5(3):187–99.

82. Paul M, Held L, Toschke AM. Multivariate modelling of infectious disease surveillance data. Stat Med. 2008;27(29):6250–67.

83. Lunn D, Thomas A, Best N, Spiegelhalter D. WinBUGS - A Bayesian Modelling Framework: Concepts, Structure, and Extensibility. Stat Comput. 2000;10:325–37.

84. Spiegelhalter D, Best N, Carlin B, van der Linde A. Bayesian Measures of Model Complexity and Fit. J R Stat Soc Ser B-Stat Methodol. 2002;64(4):583–616.

85. Kaler J, Hussain A, Patel K, Hernandez T, Ray S. Respiratory Syncytial Virus: A Comprehensive Review of Transmission, Pathophysiology, and Manifestation. Cureus. 15(3):e36342.

86. Kenmoe S, Kengne-Nde C, Ebogo-Belobo JT, Mbaga DS, Fatawou Modiyinji A, Njouom R. Systematic review and meta-analysis of the prevalence of common respiratory viruses in children < 2 years with bronchiolitis in the pre-COVID-19 pandemic era. PloS One. 2020;15(11):e0242302.

87. Dagan R, Danino D, Weinberger DM. The Pneumococcus-Respiratory Virus Connection-Unexpected Lessons From the COVID-19 Pandemic. JAMA Netw Open. 2022 Jun 1;5(6):e2218966.

88. Florin TA, Plint AC, Zorc JJ. Viral bronchiolitis. The Lancet. 2017 Jan 14;389(10065):211–24.

89. Smyth RL, Openshaw PJ. Bronchiolitis. The Lancet. 2006 Jul 22;368(9532):312–22.

90. Heppe Montero M, Gil-Prieto R, Walter S, Aleixandre Blanquer F, Gil De Miguel Á. Burden of severe bronchiolitis in children up to 2 years of age in Spain from 2012 to 2017. Hum Vaccines Immunother. 2022 Dec 31;18(1):1883379.

91. Mira-Iglesias A, Demont C, López-Labrador FX, Mengual-Chuliá B, García-Rubio J, Carballido-Fernández M, et al. Role of age and birth month in infants hospitalized with RSV-confirmed disease in the Valencia Region, Spain. Influenza Other Respir Viruses. 2022;16(2):328–39.

92. Rybak A, Levy C, Jung C, Béchet S, Batard C, Hassid F, et al. Delayed Bronchiolitis Epidemic in French Primary Care Setting Driven by Respiratory Syncytial Virus: Preliminary Data from the Oursyn Study, March 2021. Pediatr Infect Dis J. 2021 Dec;40(12):e511.

93. Sánchez Luna M, Pérez Muñuzuri A, Leante Castellanos JL, Ruiz Campillo CW, Sanz López E, Benavente Fernández I, et al. Recomendaciones de la Sociedad Española de Neonatología para la utilización de palivizumab como profilaxis de las infecciones graves por el virus respiratorio sincitial en lactantes de alto riesgo, actualización. An Pediatría. 2019 Nov;91(5):348–50.

94. Kieffer A, Beuvelet M, Sardesai A, Musci R, Milev S, Roiz J, et al. Expected Impact of Universal Immunization With Nirsevimab Against RSV-Related Outcomes and Costs Among All US Infants in Their First RSV Season: A Static Model. J Infect Dis. 2022 Aug 15;226(Supplement_2):S282–92.

95. Hammitt LL, Dagan R, Yuan Y, Baca Cots M, Bosheva M, Madhi SA, et al. Nirsevimab for Prevention of RSV in Healthy Late-Preterm and Term Infants. N Engl J Med. 2022 Mar 3;386(9):837–46.

96. Keam SJ. Nirsevimab: First Approval. Drugs. 2023 Feb;83(2):181–7.

97. Kampmann B, Madhi SA, Munjal I, Simões EAF, Pahud BA, Llapur C, et al. Bivalent Prefusion F Vaccine in Pregnancy to Prevent RSV Illness in Infants. N Engl J Med. 2023 Apr 20;388(16):1451–64.

98. López-Lacort M, Corberán-Vallet A, Santonja Gómez FJ. A Multivariate Age-Structured Stochastic Model with Immunization Strategies to Describe Bronchiolitis Dynamics. Int J Environ Res Public Health. 2021 Jan;18(14):7607.

99. Prasad N, Read JM, Jewell C, Waite B, Trenholme AA, Huang QS, et al. Modelling the impact of respiratory syncytial virus (RSV) vaccine and immunoprophylaxis strategies in New Zealand. Vaccine. 2021 Jul 13;39(31):4383–90.

100. Baral R, Li X, Willem L, Antillon M, Vilajeliu A, Jit M, et al. The impact of maternal RSV vaccine to protect infants in Gavi-supported countries: Estimates from two models. Vaccine. 2020 Jul 14;38(33):5139–47.

101. Muñoz-Quiles C, López-Lacort M, Díez-Domingo J, Orrico-Sánchez A. Bronchiolitis, regardless of its aetiology and severity, is associated with an increased risk of asthma: a population-based study. J Infect Dis. 2023 Apr 5;jiad093.

102. RSV-Pediatric-05-Jones-508.pdf [Internet]. [cited 2023 Jul 12]. Available from: https://www.cdc.gov/vaccines/acip/meetings/downloads/slides-2023-02/slides-02-23/RSV-Pediatric-05-Jones-508.pdf

103. Groupe SA. GlobeNewswire News Room. 2022 [cited 2022 Dec 19]. Press Release: New nirsevimab data analyses reinforce efficacy against RSV. Available from: https://www.globenewswire.com/fr/news-release/2022/05/11/2440425/0/en/Press-Release-New-nirsevimab-data-analyses-reinforce-efficacy-against-RSV.html

104. Simões EAF, Madhi SA, Muller WJ, Atanasova V, Bosheva M, Cabañas F, et al. Efficacy of nirsevimab against respiratory syncytial virus lower respiratory tract infections in preterm and term infants, and pharmacokinetic extrapolation to infants with congenital heart disease and chronic lung disease: a pooled analysis of randomised controlled trials. Lancet Child Adolesc Health. 2023 Mar;7(3):180–9.

105. Wilkins D, Yuan Y, Chang Y, Aksyuk AA, Núñez BS, Wählby-Hamrén U, et al. Durability of neutralizing RSV antibodies following nirsevimab administration and elicitation of the natural immune response to RSV infection in infants. Nat Med. 2023 May;29(5):1172–9.

106. Economics of preventing respiratory syncytial virus lower respiratory tract infections (RSVLRTI) among US infants with Nirsevimab : a summary report comparing models from: Sanofi and University of Michigan and CDC [Internet]. [cited 2023 Aug 2]. Available from: https://stacks.cdc.gov/view/cdc/125143

107. Economics of Pfizer maternal RSVpreF vaccine [Internet]. [cited 2023 Aug 2]. Available from: https://stacks.cdc.gov/view/cdc/130019

108. Munjal I. Safety and Efficacy of Bivalent RSV Prefusion F Vaccine in Vaccinated Mothers and their Infants. Worldw Res. 2018;

109. Todas_las_tablas2020.pdf [Internet]. [cited 2023 Sep 6]. Available from: https://www.sanidad.gob.es/areas/promocionPrevencion/vacunaciones/calendario-y-coberturas/coberturas/docs/Todas_las_tablas2020.pdf

110. Simon Drysdale S, Cathie K , Flamein F, Knuf M, Collins A, Hill H, Kaiser F et al.: Efficacy of nirsevimab against RSV lower respiratory tract infection hospitalization in infants: preliminary data from the Harmonie phase 3B trial. Abstract book ESPID 2023.

111. anx_160227_es.pdf [Internet]. [cited 2023 Sep 26]. Available from: https://ec.europa.eu/health/documents/community-register/2023/20230823160227/anx_160227_es.pdf

112. Pfizer. A Phase3, randomized, double-blinded, placebo-controlled trial to evaluate the efficacy and safety of a respiratory syncytial virus (RSV) prefusion F subunit vaccine in infants born to women vaccinated during pregnancy [Internet]. clinicaltrials.gov; 2023 Jun [cited 2023 Jul 12]. Report No.: NCT04424316. Available from: https://clinicaltrials.gov/study/NCT04424316

113. Merali Z, Wilson JR. Explanatory Versus Pragmatic Trials: An Essential Concept in Study Design and Interpretation. Clin Spine Surg. 2017 Nov;30(9):404–6.

114. Sanofi Pasteur, a Sanofi Company. A Phase IIIb Randomized Open-label Study of Nirsevimab (Versus no Intervention) in Preventing Hospitalizations Due to Respiratory Syncytial Virus in Infants (HARMONIE) [Internet]. clinicaltrials.gov; 2023 Aug [cited 2023 Sep 5]. Report No.: NCT05437510. Available from: https://clinicaltrials.gov/study/NCT05437510

115. Hartwig S, Baldauf JJ, Dominiak-Felden G, Simondon F, Alemany L, de Sanjosé S, et al. Estimation of the epidemiological burden of HPV-related anogenital cancers, precancerous lesions, and genital warts in women and men in Europe: Potential additional benefit of a nine-valent second generation HPV vaccine compared to first generation HPV vaccines. Papillomavirus Res. 2015;1:90–100.

116. FUTURE II Study Group. Quadrivalent vaccine against human papillomavirus to prevent high-grade cervical lesions. N Engl J Med. 2007;356:1915–27.

117. Paavonen J, Naud P, Salmerón J, Wheeler CM, Chow SN, Apter D, et al. Efficacy of human papillomavirus (HPV)-16/18 AS04-adjuvanted vaccine against cervical infection and precancer caused by oncogenic HPV types (PATRICIA): final analysis of a double-blind, randomised study in young women. N Engl J Med. 2009;374:301–14.

118. Arbyn M, Xu L, Simoens C, Martin-Hirsch PP. Prophylactic vaccination against human papillomaviruses to prevent cervical cancer and its precursors. Cochrane Database Syst Rev. 2018;5:CD009069.

119. Leval A, Herweijer E, Ploner A, Eloranta S, Fridman Simard J, Dillner J, et al. Quadrivalent human papillomavirus vaccine effectiveness: a Swedish national cohort study. J Natl Cancer Inst. 2013;105:469–74.

120. Garland SM, Kjaer SK, Muñoz N, Block SL, Brown DR, DiNubile MJ, et al. Impact and effectiveness of the quadrivalent human papillomavirus vaccine: A systematic review of 10 years of real-world experience. Clin Infect Dis. 2016;63:519–27.

121. Silverberg MJ, Leyden WA, Lam JO, Gregorich SE, Huchko MJ, Kulasingam S, et al. Effectiveness of catch-up human papillomavirus vaccination on incident cervical neoplasia in a US health-care setting: a population-based case-control study. Lancet Child Adolesc Health. 2018;2:707–14.

122. Drolet M, Bénard É, Pérez N, Brisson M, Group HVIS. Population-level impact and herd effects following the introduction of human papillomavirus vaccination programmes: updated systematic review and meta-analysis. N Engl J Med. 2019;394:497–509.

123. Lei J, Ploner A, Elfstr\”om KM, Wang J, Roth A, Fang F, et al. HPV vaccination and the risk of invasive cervical cancer. N Engl J Med. 2020;383:1340–8.

124. Kjaer SK, Dehlendorff C, Belmonte F, Baandrup L. Real-world effectiveness of human papillomavirus vaccination against cervical cancer. J Natl Cancer Inst. 2021;113:1329–35.

125. Ministerio de Sanidad. Portal estadístico. Área de Inteligencia de Gestión [Internet]. Available from: https://pestadistico.inteligenciadegestion.sanidad.gob.es/publicoSNS/I/sivamin/informe-de-evolucion-de-coberturas-de-vacunacion-por-vacuna, year = 2023

126. Navarro-Illana P, Caballero P, Tuells J, Puig-Barberá J, Diez-Domingo J. Aceptabilidad de la vacuna contra el virus del papiloma humano en madres de la provincia de Valencia (España). An Pediatría. 2015 Nov 1;83(5):318–27.

127. Finney Rutten LJ, Wilson PM, Jacobson DJ, Agunwamba AA, Radecki Breitkopf C, Jacobson RM, et al. A Population-Based Study of Sociodemographic and Geographic Variation in HPV Vaccination. Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol. 2017 Apr;26(4):533–40.

128. Riesen M, Konstantinoudis G, Lang P, Low N, Hatz C, Maeusezahl M, et al. Exploring variation in human papillomavirus vaccination uptake in Switzerland: a multilevel spatial analysis of a national vaccination coverage survey. Open Access.

129. Wheeler DC, Miller CA, Do EK, Ksinan AJ, Trogdon JG, Chukmaitov A, et al. Identifying Area-Level Disparities in Human Papillomavirus Vaccination Coverage Using Geospatial Analysis. Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol. 2021 Sep;30(9):1689–96.

130. Pourebrahim N, Shah P, VoPham T, Doody DR, Bell TR, deHart MP, et al. Time and geographic variations in human papillomavirus vaccine uptake in Washington state. Prev Med. 2021 Dec 1;153:106753.

131. Lee D, Rushworth A, Napier G. Spatio-temporal areal unit modeling in R with conditional autoregressive priors using the CARBayesST package. J Stat Softw. 2018;84:1–39.

132. Lawson AB. Bayesian disease mapping: Hierarchical modeling in spatial epidemiology, 3rd Edition. Chapman & Hall/CRC; 2018.

133. Bernardinelli L, Clayton D, Pascutto C, Montomoli C, Ghislandi M, Songini M. Bayesian analysis of space-time variation in disease risk. Stat Med. 1995;14:2433–43.

134. Knorr-Held L. Bayesian modelling of inseparable space-time variation in disease risk. Stat Med. 2000;19:2555–67.

135. Adin A, Lee D, Goicoa T, Ugarte MD. A two-stage approach to estimate spatial and spatio-temporal disease risks in the presence of local discontinuities and clusters. Stat Methods Med Res. 2019;28:2595–613.

136. Knorr-Held L, Rasser G. Bayesian detection of clusters and discontinuities in disease maps. Biometrics. 2000;56:13–21.

137. Denison DGT, Holmes CC. Bayesian partitioning for estimating disease risk. Biometrics. 2001;57:143–9.

138. Green PJ, Richardson S. Hidden Markov models and disease mapping. J Am Stat Assoc. 2002;97:1055–70.

139. Charras-Garrido M, Abrial D, De Goer J, Dachian S, Peyrard N. Classification method for disease risk mapping based on discrete hidden Markov random fields. Biostatistics. 2012;13:241–55.

140. Gómez-Rubio V, Moraga P, Molitor J, Rowlingson B. DClusterm: Model-based detection of disease clusters. J Stat Softw. 2019;90(14):1–26.

141. Kulldorff M. A spatial scan statistic. Commun Stat - Theory Methods. 1997;26:1481–96.

142. Jaya IGNM, Folmer H. Identifying spatiotemporal clusters by means of agglomerative hierarchical clustering and Bayesian regression analysis with spatiotemporally varying coefficients: methodology and application to dengue disease in Bandung, Indonesia. Geogr Anal. 2021;53:767–817.

143. Kamenetsky ME, Lee J, Zhu J, Gangnon RE. Regularized spatial and spatio-temporal cluster detection. Spat Spatio-Temporal Epidemiol. 2022;4:100462.

144. Napier G, Lee D, Robertson C, Lawson A. A Bayesian space-time model for clustering areal units based on their disease trends. Biostatistics. 2019;20:681–97.

145. Mozdzen A, Cremaschi A, Cadonna A, Guglielmi A, Kastner G. Bayesian modeling and clustering for spatio-temporal areal data: an application to Italian unemployment. Spat Stat. 2022;52:100715.

146. Nasserinejad K, van Rosmalen J, de Kort W, Lesaffre E. Comparison of criteria for choosing the number of classes in Bayesian finite mixture models. PLoS ONE. 2017;12(1):e0168838.

147. Watanabe S. Asymptotic equivalence of the Bayes cross validation and widely applicable information criterion in singular learning theory. J Mach Learn Res. 2010;11:3571–94.

148. de Valpine P, Turek D, Paciorek CJ, Anderson-Bergman C, Lang DT, Bodik R. Programming with models: writing statistical algorithms for general model structures with NIMBLE. J Comput Graph Stat. 2017;26:403–13.

149. R Core Team. R: A language and environment for statistical computing [Internet]. R Foundation for Statistical Computing; 2023. Available from: https://www.R-project.org/

150. Wickham H. ggplot2: Elegant graphics for data analysis. Springer-Verlag, New York; 2016.

151. Bivand R, Keitt T, Rowlingson B. rgdal: Bindings for the `geospatial' data abstraction library [Internet]. 2019. Available from: https://CRAN.R-project.org/package=rgdal

152. Bivand R, Rundel C. rgeos: Interface to geometry engine - open source ('GEOS') [Internet]. 2020. Available from: https://CRAN.R-project.org/package=rgeos

153. Bivand RS, Wong DWS. Comparing implementations of global and local indicators of spatial association. TEST. 2018;27:716–48.

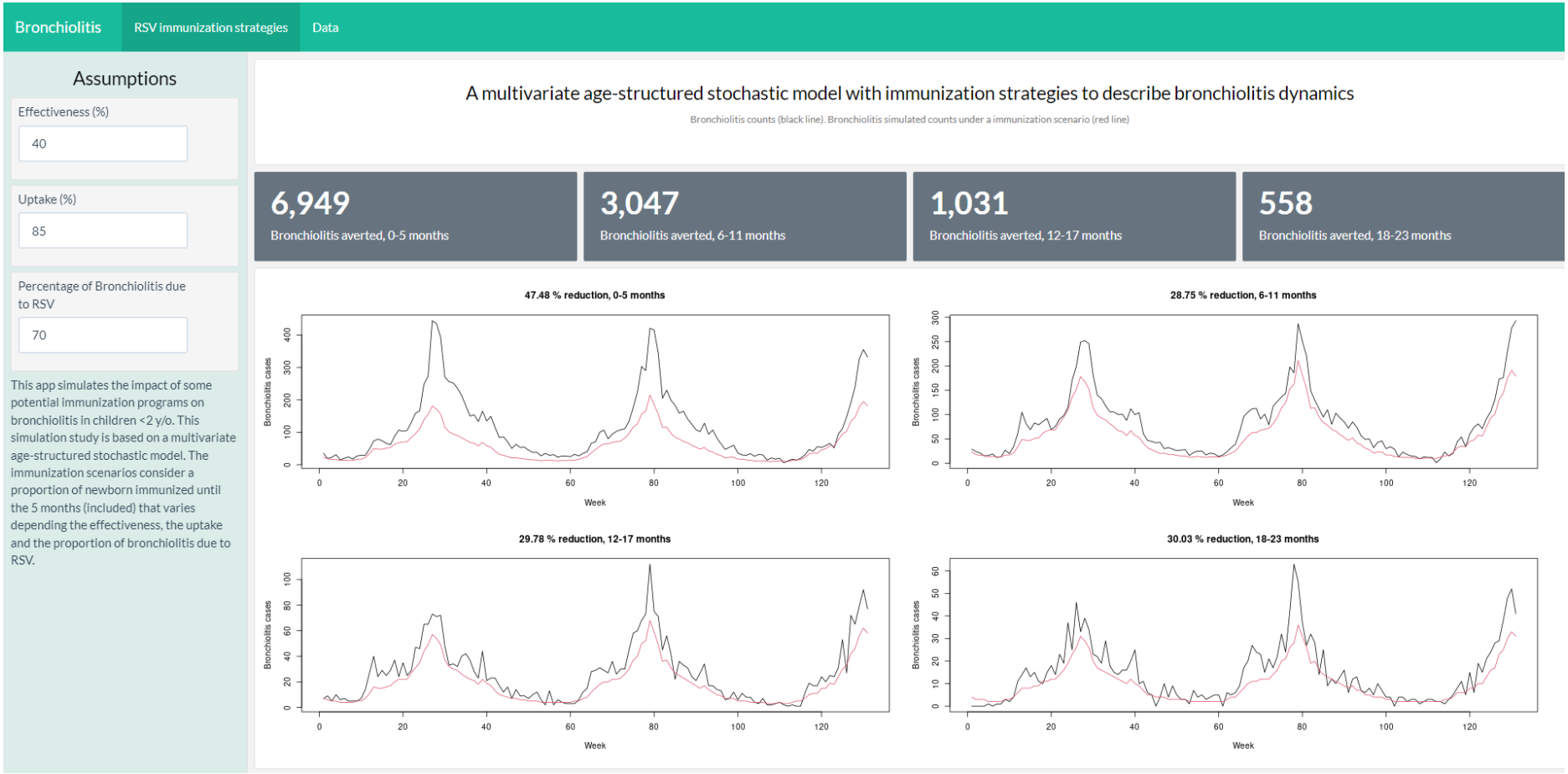
Annexes

Annex 2.1: WinBUGS code

```
#-------------------------------------------------------
#                  PROPOSED MODEL
#-------------------------------------------------------
   model{
   pi<-3.1416

   for(j in 1:4)
   {
      y[1,j] ~ dpois(mu[1,j])
      mu[1,j]<-exp(lambda[1,j])*pow(sum(y0[1:4]),h)
      lambda[1,j]<- a[j]+b1*sin(2*pi/52)+g1*cos(2*pi/52)+eps[1]
   }
   for(i in 2:T){
   for(j in 1:4)
   {
       y[i,j] ~ dpois(mu[i,j])
       mu[i,j]<-exp(lambda[i,j])*pow(sum(y[i-1,]),h)
       lambda[i,j]<- a[j]+b1*sin(2*pi*t[i]/52)+g1*cos(2*pi*t[i]/52)+eps[i]
    }}

    # Priors
    h ~ dunif(0,1)

    for(j in 1:4)
    {
       a[j]~dflat()
    }

    b1~dflat()
    g1~dflat()

    for(i in 1:T){
      eps[i]~dnorm(0,tau_eps)
    }

    tau_eps<-pow(sigma_eps,-2)
    sigma_eps~dunif(0,2)
}
```

```
#-------------------------------------------------------------
#     PROPOSED MODEL WITH NEGATIVE BINOMIAL DISTRIBUTION
#-------------------------------------------------------------
   model{
   pi<-3.1416

   for(j in 1:4)
   {
      y[1,j] ~ dnegbin(p[1,j],k)
      p[1,j]<-k/(k+mu[1,j])
      mu[1,j]<-exp(lambda[1,j])*pow(sum(y0[1:4]),h)
      lambda[1,j]<- a[j]+b1*sin(2*pi/52)+g1*cos(2*pi/52)+eps[1]
   }

   for(i in 2:T){
   for(j in 1:4)
   {
       y[i,j] ~ dnegbin(p[i,j],k)
       p[i,j]<-k/(k+mu[i,j])
       mu[i,j]<-exp(lambda[i,j])*pow(sum(y[i-1,]),h)
       lambda[i,j]<- a[j]+b1*sin(2*pi*t[i]/52)+g1*cos(2*pi*t[i]/52)+eps[i]
    }}

    # Priors
    k ~ dgamma(1,0.01)

    h ~ dunif(0,1)

    for(j in 1:4)
    {
       a[j]~dflat()
    }

    b1~dflat()
    g1~dflat()

    for(i in 1:T){
      eps[i]~dnorm(0,tau_eps)
    }
    tau_eps<-pow(sigma_eps,-2)
    sigma_eps~dunif(0,2)
    }
```

Annex 2.2: App for simulating the impact of immunization scenarios on bronchiolitis

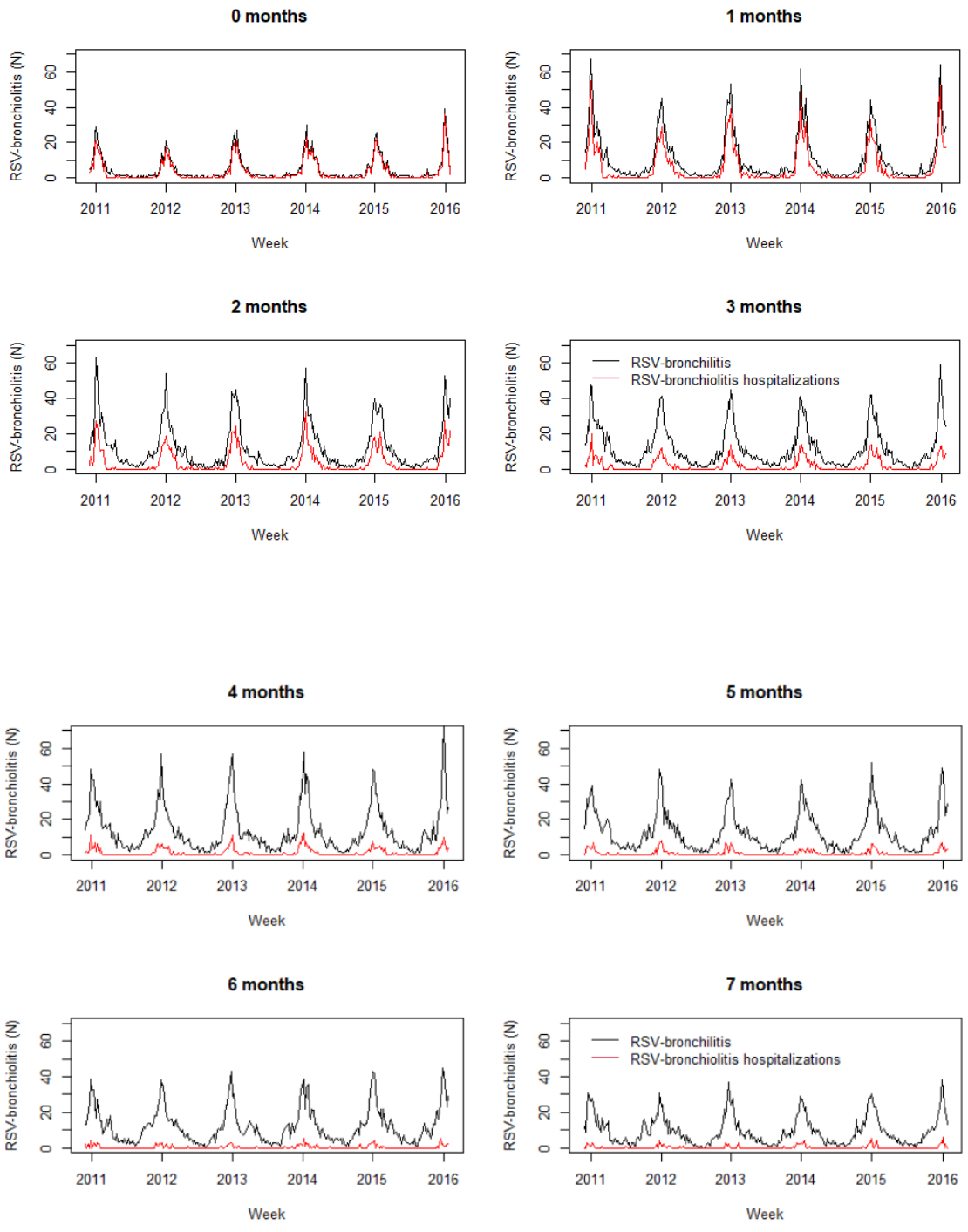


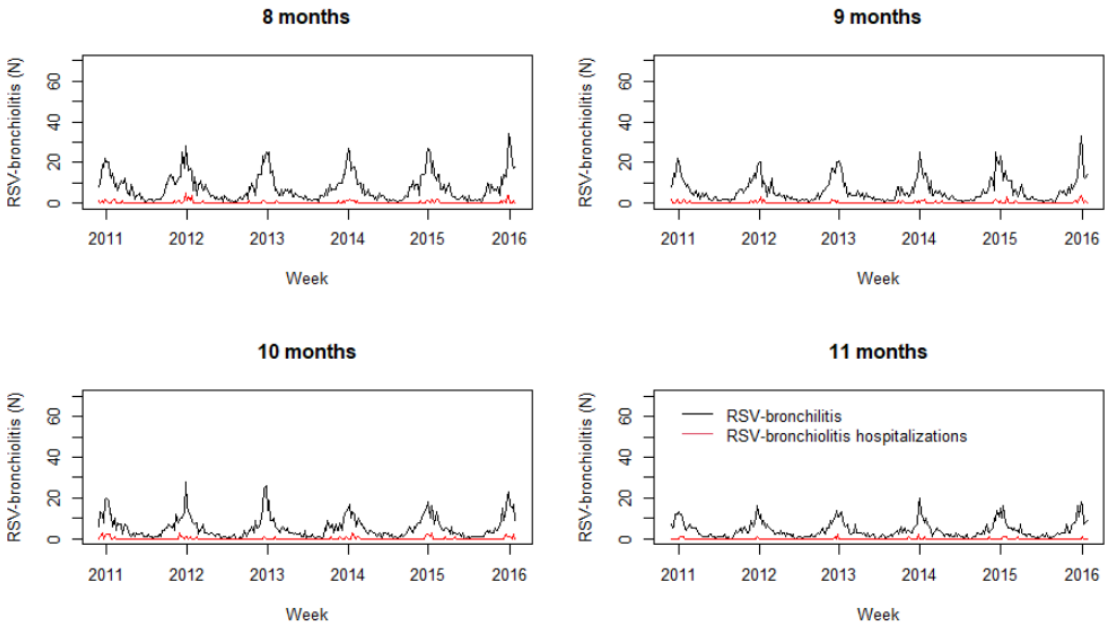
Chapter 2 app screenshot: This application simulates the potential impact of immunization programs on bronchiolitis in children < 2 y/o. By utilizing posterior parameter distributions estimated by the model, it enables us to reproduce bronchiolitis counts under different scenarios where a proportion of newborns (< 6 months of age) are immunized (red line). Black line represents count without any immunization program. Users can adjust parameters such as effectiveness, uptake, and the percentage of bronchiolitis cases due to RSV to observe the corresponding reduction in bronchiolitis cases.

Annex 3.1: WinBUGS code

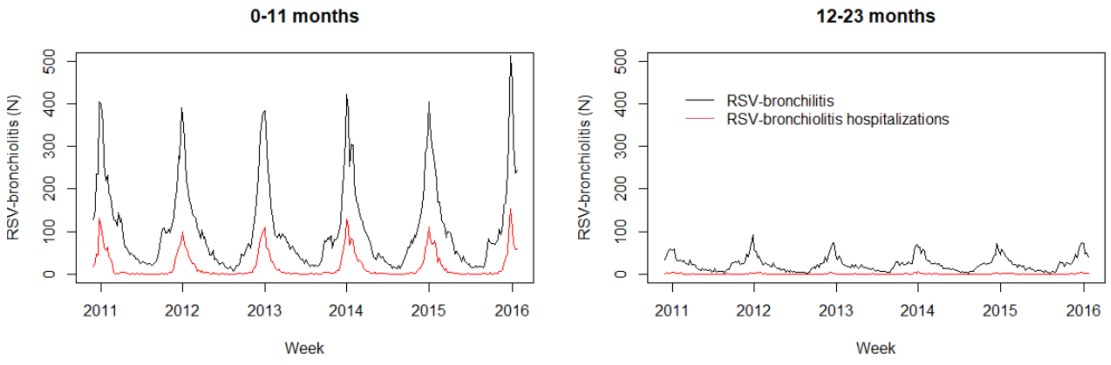
```
#-----------------------------------------------------
#                  PROPOSED MODEL
#-----------------------------------------------------

model{
pi<-3.1416

#  T=1
for(j in 1:13){
  y[1,j] ~ dbin(p[1,j],Pob[1,j])
  logit(p[1,j]) <- alpha* log(sum(Cini[1:13])) + r[1,j]}

for(k in 2:T){
for(j in 1:13){
  y[k,j] ~ dbin(p[k,j],Pob[k,j])
  logit(p[k,j]) <- alpha * log(sum(yb[k-1,])) + r[k,j]}
  }

# Model for the transmission rate using sine-cosine waves
for(k in 1:T){
for(j in 1:13){
   r[k,j] <- gama0[j] + gama1*sin(2*pi*t[k]/52) + gama2*cos(2*pi*t[k]/52) +
epsilon[k]}
epsilon[k] ~ dnorm(0,taueps)}

 # Priors for the parameters gama and the precision
  for(j in 1:13){
  gama0[j] ~ dflat()}

  gama1 ~ dflat()
  gama2 ~ dflat()

taueps <- pow(sdeps,-2)
sdeps ~ dunif(0,2)

  # Priors for the age-mixing parameters
alpha ~ dunif(0,1)
}
```

Annex 3.2: Distribution of RSV-bronchiolitis episodes by week and age group





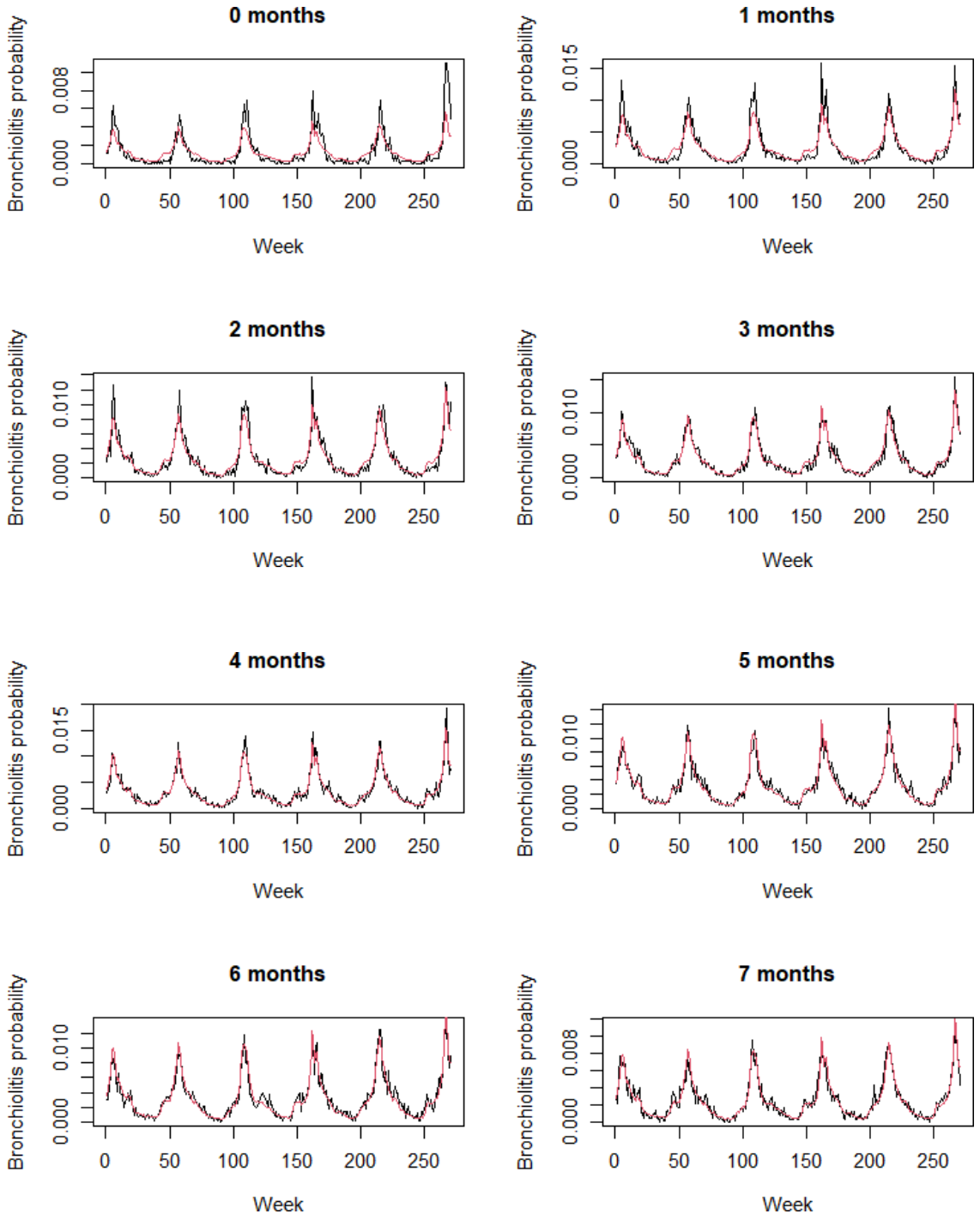
Annex 3.2a: Weekly counts (N) of RSV-bronchiolitis (solid black line) and RSV-bronchiolitis hospitalizations (red line) by age-groups up to 12 months of age from November 2010 (week 48, 2010) to January 2016 (week 4, 2016).

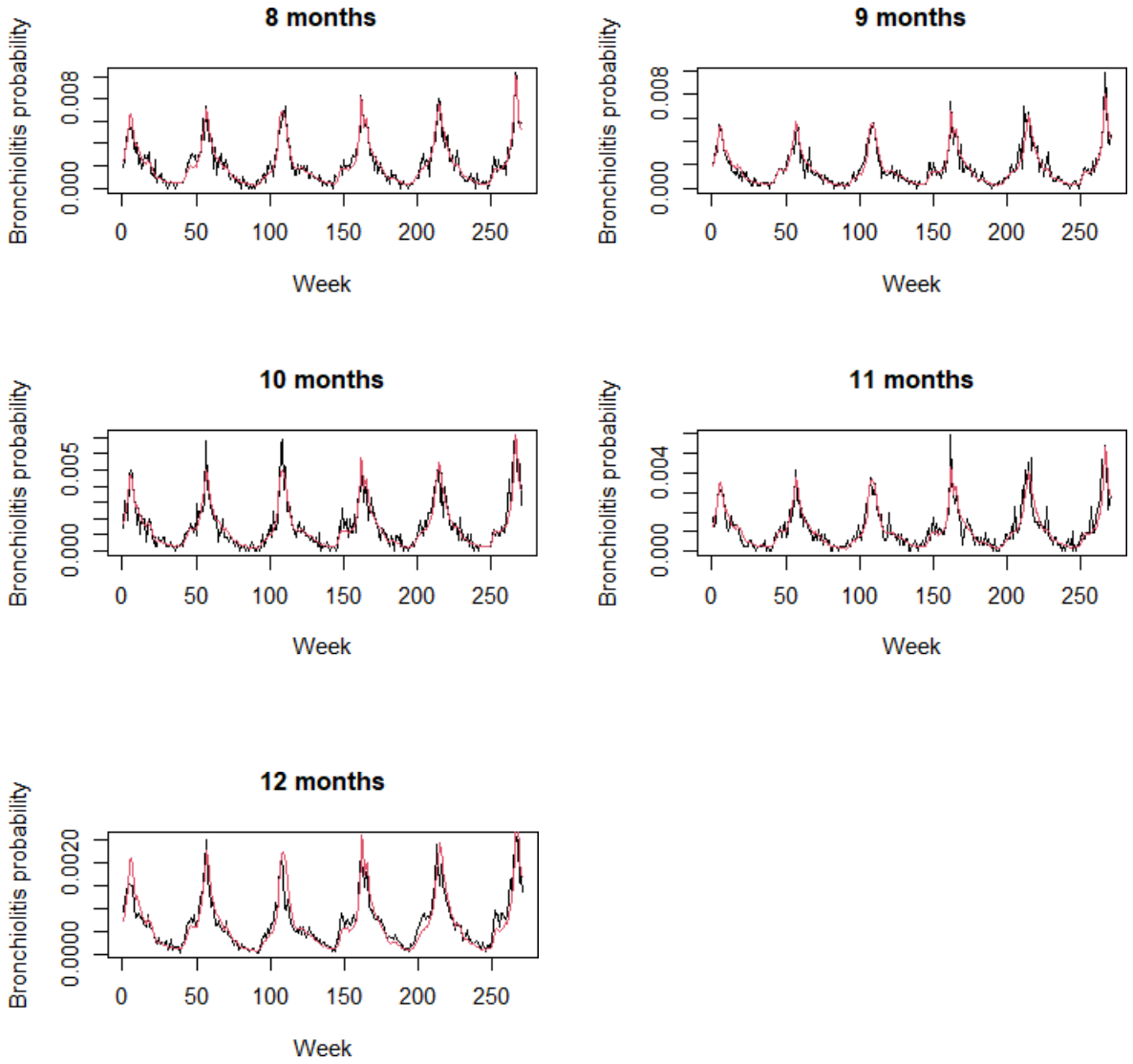


Annex 3.2b: Weekly counts (N) of RSV-bronchiolitis (solid black line) and RSV-bronchiolitis hospitalizations (red line) in children 0-11 months (which are RSV-naïve of target of immunization) and in children 12-23 months age from November 2010 (week 48, 2010) to January 2016 (week 4, 2016).

Annex 3.3: Model accuracy

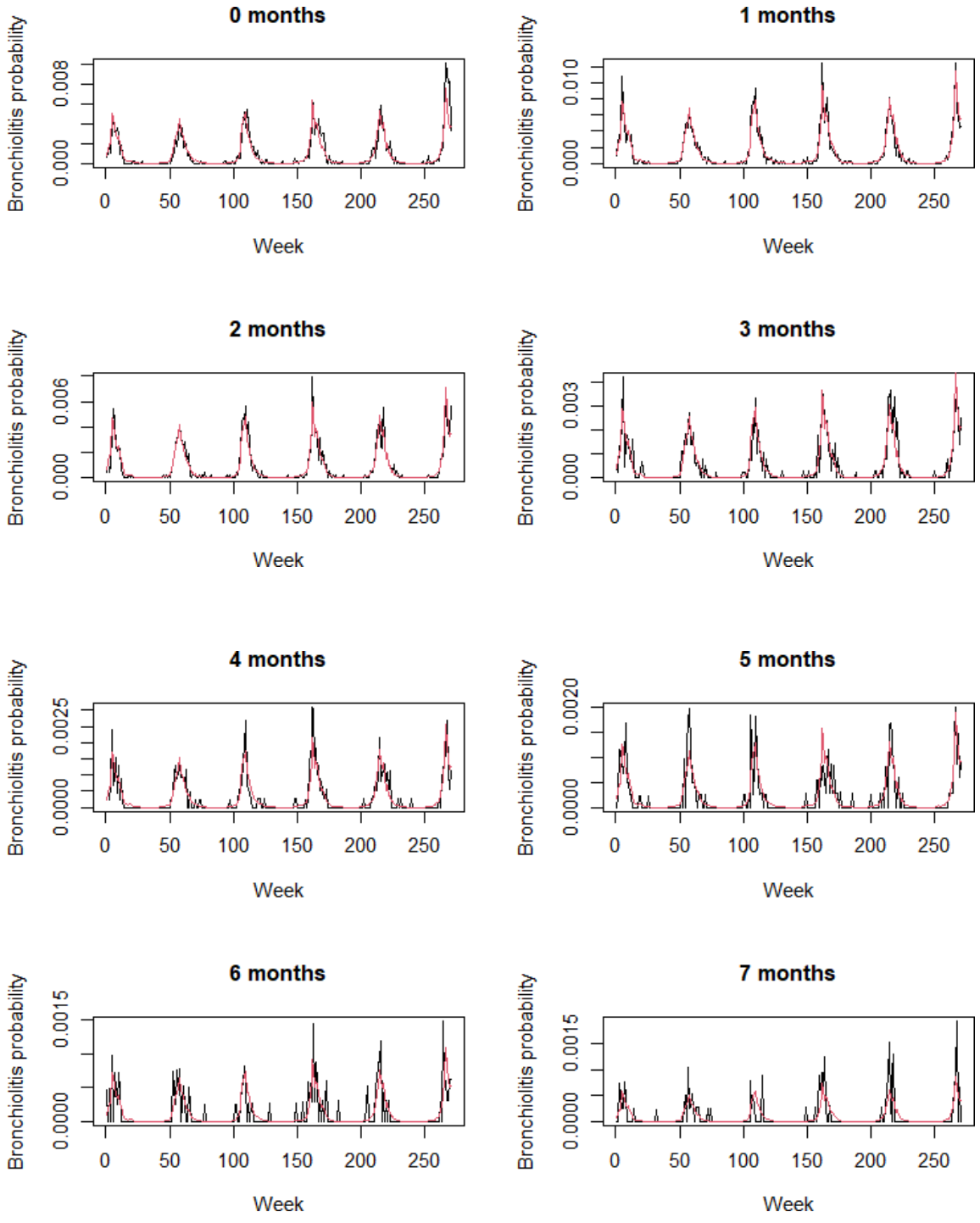
Model accuracy for RSV-bronchiolitis

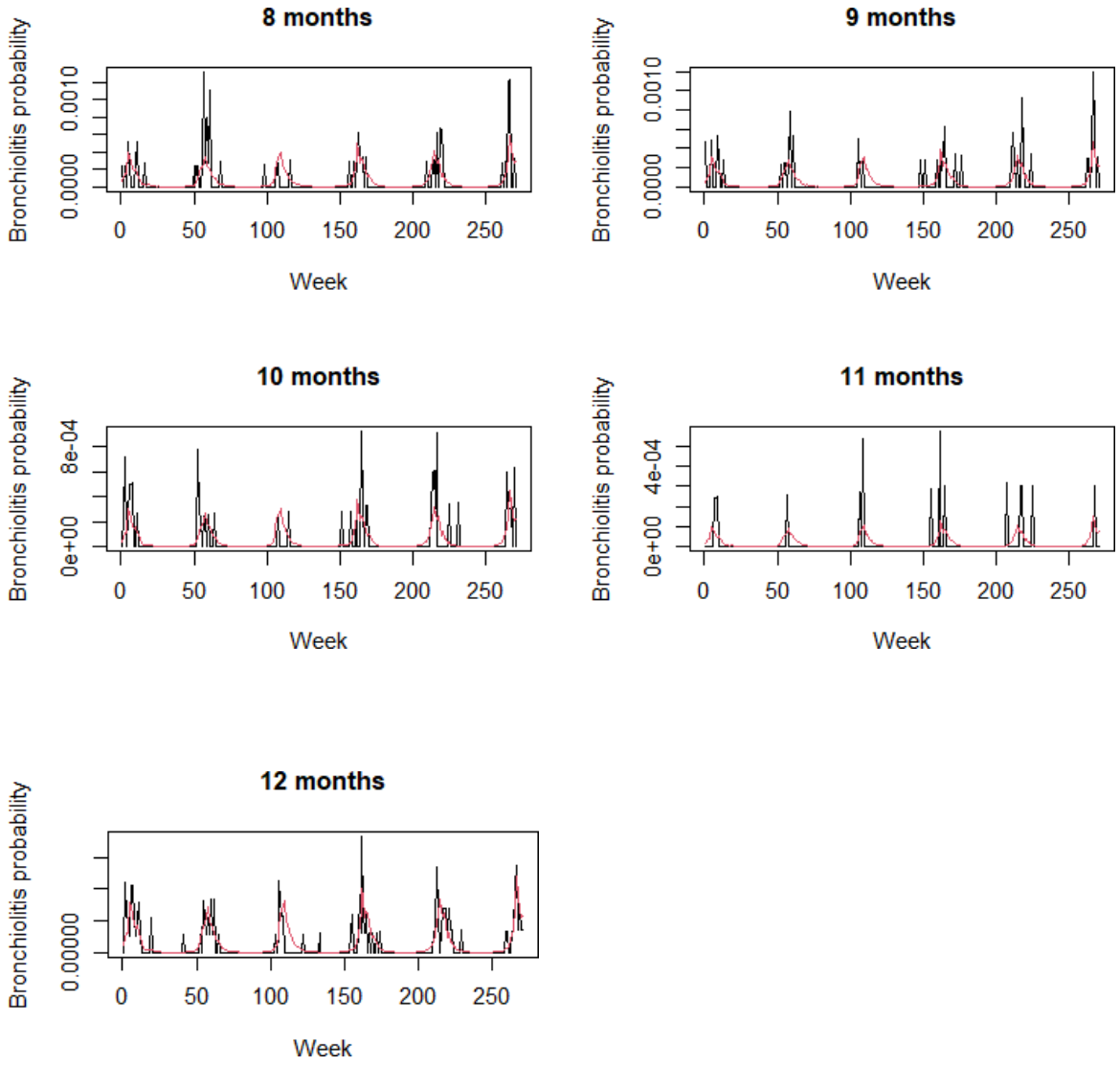




Annex 3.3a: Model accuracy, RSV-bronchiolitis observed probability (black line)together with posterior mean of the probability estimated by the model by age-group from November 2010 (week 48, 2010) to January 2016 (week 4, 2016); T=270 weeks.

Model accuracy for RSV-bronchiolitis hospitalizations





Annex 3.3b: Model accuracy, RSV-bronchiolitis hospitalizations observed probability (black line) together with posterior mean of the probability estimated by the model by age-group from November 2010 (week 48, 2010) to January 2016 (week 4, 2016); T=270 wee

Annex 3.4: Sensitivity analysis of the effectiveness

1. Assumptions

NmAb		
Sensitivity 1: -10% of main analysis effectiveness	**RSV-Bronchiolitis hospitalizations**	**RSV-Bronchiolitis**
Months of protection	Effectiveness	
0,1,2,3,4	67,3	69,5
Decay	*y= -11,22*x+112,18*	*y=-11,58*x+115,82*
5	56,1	57,9
6	44,9	46,3
7	33,6	34,8
8	22,4	23,2
9	11,2	11,6
10	0,0	0,0
Sensitivity 2: -20% of main analysis effectiveness		
0,1,2,3,4	57,3	59,5
Decay	*y=9,55*x+95,5*	*y=-9,92*x+99,18*
5	47,8	49,6
6	38,2	39,7
7	28,7	29,7
8	19,1	19,8
9	9,6	9,9
10	0,0	0,0
MI		
Sensitivity 1: -10% of main analysis effectiveness	**RSV-Bronchiolitis hospitalizations**	**RSV-Bronchiolitis**
Months of protection	Effectiveness	
0,1,2,3,4,5	46,8	41,3
Decay	*y=9,36*x+93,6*	*y= -8,26*x+82,6*
6	37,4	33,0
7	28,1	24,8
8	18,7	16,5
9	9,4	8,3
10	0,0	0,0
Sensitivity 2: -10% of main analysis effectiveness		
0,1,2,3,4,5	36,8	31,3
Decay	*y=-7,36*x+73,6*	*y= -6,26*x+62,6*
6	29,4	25,0
7	22,1	18,8
8	14,7	12,5
9	7,4	6,3
10	0,0	0,0

Annex 3.4_table_1: Impact simulation assumptions for the effectiveness sensitivity analysis: Sensitivity 1 and 2 assume a 10% and 20% lower effectiveness, respectively than the base case for the first 5 months of protection for NmAb and the first 6 months for MI, followed by a linear decay to 0.Equations represent the decay were, x, are the specific month after immunization and, y, the estimated effectiveness.

2. Sensitivity 1: Impact simulation of Immunization Strategies

A) NmAb Impact estimations (Nov 10 – Jan 16)

	RSV-bronchiolitis 29872 (29533, 30198) *				
	N (95% CrI) averted bronchiolitis; impact %				
Program/Uptake	**55%**	**65%**	**75%**	**85%**	**95%**
Catch-up & Seasonal	9653 (9538,9759) 32%	11408 (11272,11534) 38%	13163 (13006,13308) 44%	14918 (14740,15082) 50%	16673 (16474,16857) 56%
	RSV-bronchiolitis hospitalizations 4924 (4794, 5072) *				
Catch-up & Seasonal	1767 (1720,1819) 36%	2088 (2032,2150) 42%	2410 (2345,2481) 49%	2731 (2657,2812) 55%	3052 (2970,3142) 62%

B) NmAb Impact estimations annually from Oct -Sep **

	RSV-bronchiolitis 15561(15215,15909) *				
	N (95% CrI) averted bronchiolitis per 100,00 children-yr; impact %				
Program/Uptake	**55%**	**65%**	**75%**	**85%**	**95%**
Catch-up & Seasonal	4992 (4875,5107) 32%	5899 (5761,6035) 38%	6807 (6648,6964) 44%	7714 (7534,7893) 50%	8622 (8420,8821) 55%
	RSV-bronchiolitis hospitalizations 2482(2336,2642) *				
Catch-up & Seasonal	888 (834,947) 36%	1050 (986,1119) 42%	1211 (1137,1291) 49%	1373 (1289,1463) 55%	1535 (1440,1636) 62%

Annex 3.4_table_2: Impact of seasonal with catch-up NmAb program on RSV-bronchiolitis and RSV-bronchiolitis hospitalizations in sensitivity analysis 1. A) Provides the average number of bronchiolitis cases averted with the corresponding impact percentages, for the whole study period and B) by 100,000 infants per year. Different uptake MI levels (55% to 95%) were assessed. 95%

*CrI: 95% credible interval. * Number of predicted bronchiolitis without immunization program. ** The annual (October-September) average of number of averted bronchiolitis per 100,000 children was estimated excluding the incomplete periods 1) 2010-Nov, 2011-Sep and 2) 2015-Oct, 2016-Jan.*

A) MI Impact estimations (Nov 10 – Jan 16)

RSV-bronchiolitis 29872 (29533, 30198) *					
	N (95% CrI) averted bronchiolitis; impact %				
Program/Uptake	**55%**	**65%**	**75%**	**85%**	**95%**
Seasonal	2057 (2026,2086) 7%	2432 (2394,2465) 8%	2806 (2763,2845) 9%	3180 (3131,3224) 11%	3554 (3500,3603) 12%
Year-round	6198 (6120,6266) 21%	7324 (7232,7405) 25%	8451 (8345,8544) 28%	9578 (9458,9683) 32%	10705 (10571,10823) 36%
RSV-bronchiolitis hospitalizations 4924 (4794, 5072) *					
Seasonal	793 (767,820) 16%	937 (907,969) 19%	1081 (1046,1118) 22%	1226 (1185,1267) 25%	1370 (1325,1416) 28%
Year-round	1250 (1217,1287) 25%	1477 (1438,1521) 30%	1705 (1660,1756) 35%	1932 (1881,1990) 39%	2159 (2102,2224) 44%

**B) MI Impact estimations annually from Oct -Sep **

RSV-bronchiolitis 15561(15215,15909) *					
	N (95% CrI) averted bronchiolitis per 100,000 children-yr; impact %				
Program/Uptake	**55%**	**65%**	**75%**	**85%**	**95%**
Seasonal	1096 (1069,1124) 7%	1296 (1263,1328) 8%	1495 (1458,1532) 10%	1694 (1652,1737) 11%	1893 (1846,1941) 12%
Year-round	3228 (3154,3301) 21%	3814 (3727,3901) 25%	4401 (4301,4501) 28%	4988 (4874,5101) 32%	5575 (5448,5702) 36%
RSV-bronchiolitis hospitalizations 2482(2336,2642) *					
Seasonal	406 (380,433) 16%	479 (449,512) 19%	553 (518,591) 22%	627 (587,669) 25%	700 (656,748) 28%
Year-round	630 (593,671) 25%	745 (700,793) 30%	859 (808,915) 35%	974 (916,1037) 39%	1088 (1024,1159) 44%

*Annex 3.4_table_3: Impact of seasonal and year-round maternal immunization programs on RSV-bronchiolitis and RSV-bronchiolitis hospitalizations in sensitivity analysis 1. A) Provides the average number of bronchiolitis cases averted with the corresponding impact percentages, for the whole study period and B) by 100,000 infants per year. Different uptake MI levels (55% to 95%) were assessed. 95% CrI: 95% credible interval. * Number of predicted bronchiolitis without immunization program. ** The annual (October-September) average of number of averted bronchiolitis per 100,000 children was estimated excluding the incomplete periods 1) 2010-Nov, 2011-Sep and 2) 2015-Oct, 2016-Jan.*

3. Sensitivity 2: Impact simulation of Immunization Strategies

A) NmAb Impact estimations (Nov 10 – Jan 16)

RSV-bronchiolitis 29872 (29533, 30198) *					
	N (95% CrI) averted bronchiolitis; impact %				
Program/Uptake	**55%**	**65%**	**75%**	**85%**	**95%**
Catch-up & Seasonal	8264 (8165,8355) 28%	9767 (9650,9874) 33%	11269 (11135,11393) 38%	12772 (12619,12912) 43%	14274 (14104,14431) 48%
RSV-bronchiolitis hospitalizations 4924 (4794, 5072) *					
Catch-up & Seasonal	1504 (1464,1549) 31%	1778 (1730,1831) 36%	2052 (1996,2112) 42%	2325 (2262,2394) 47%	2599 (2529,2675) 53%

B) NmAb Impact estimations annually from Oct -Sep **

RSV-bronchiolitis 15561(15215,15909) *					
	N (95% CrI) averted bronchiolitis per 100,00 children-yr; impact %				
Program/Uptake	**55%**	**65%**	**75%**	**85%**	**95%**
Catch-up & Seasonal	4273 (4174,4372) 27%	5050 (4932,5167) 32%	5827 (5691,5962) 37%	6604 (6450,6757) 42%	7381 (7209,7552) 47%
RSV-bronchiolitis hospitalizations 2482(2336,2642) *					
Catch-up & Seasonal	756 (710,806) 30%	894 (839,953) 36%	1031 (968,1099) 42%	1169 (1097,1246) 47%	1307 (1226,1392) 53%

*Annex 3.4_table_4: Impact of seasonal with catch-up NmAb program on RSV-bronchiolitis and RSV-bronchiolitis hospitalizations in sensitivity analysis 2. A) Provides the average number of bronchiolitis cases averted with the corresponding impact percentages, for the whole study period and B) by 100,000 infants per year. Different uptake MI levels (55% to 95%) were assessed. 95% CrI: 95% credible interval. * Number of predicted bronchiolitis without immunization program. ** The annual (October-September) average of number of averted bronchiolitis per 100,000 children was estimated excluding the incomplete periods 1) 2010-Nov, 2011-Sep and 2) 2015-Oct, 2016-Jan.*

A) MI Impact estimations (Nov 10 – Jan 16)

RSV-bronchiolitis 29872 (29533, 30198) *					
	N (95% CrI) averted bronchiolitis; impact %				
Program/Uptake	**55%**	**65%**	**75%**	**85%**	**95%**
Seasonal	1559 (1536,1581) 5%	1843 (1815,1868) 6%	2126 (2094,2156) 7%	2410 (2373,2443) 8%	2693 (2652,2731) 9%
Year-round	4697 (4638,4749) 16%	5551 (5481,5612) 19%	6405 (6325,6475) 21%	7259 (7168,7339) 24%	8113 (8011,8202) 27%
RSV-bronchiolitis hospitalizations 4924 (4794, 5072) *					
Seasonal	624 (603,644) 13%	737 (713,762) 15%	850 (822,879) 17%	964 (932,996) 20%	1077 (1042,1113) 22%
Year-round	983 (957,1012) 20%	1162 (1131,1196) 24%	1340 (1305,1380) 27%	1519 (1479,1564) 31%	1698 (1653,1749) 34%

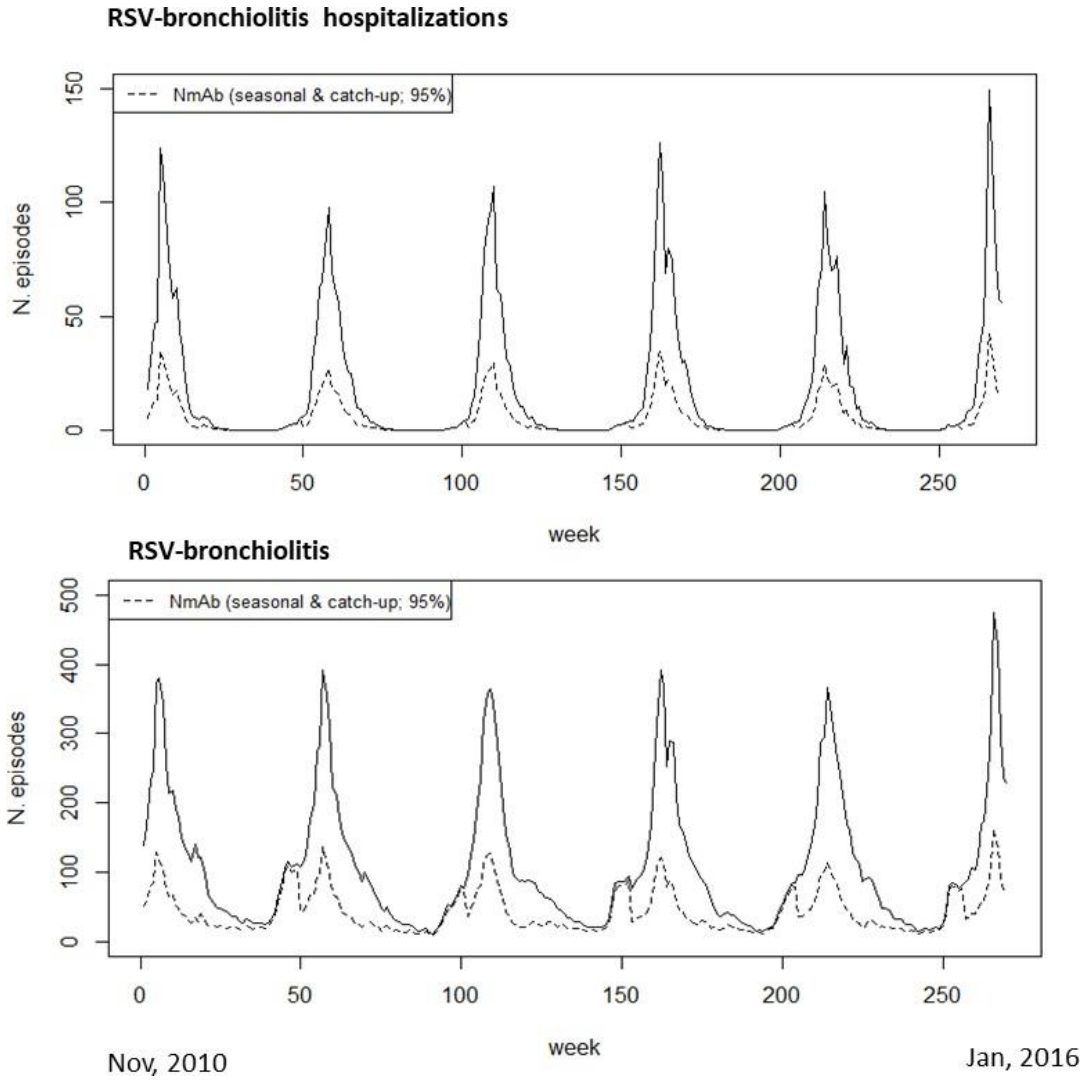
B) MI Impact estimations annually from Oct -Sep **

RSV-bronchiolitis 15561(15215,15909) *					
	N (95% CrI) averted bronchiolitis per 100,000 children-yr; impact %				
Program/Uptake	**55%**	**65%**	**75%**	**85%**	**95%**
Seasonal	831 (810,852) 5%	982 (957,1006) 6%	1133 (1105,1161) 7%	1284 (1252,1316) 8%	1435 (1399,1471) 9%
Year-round	2446 (2390,2502) 16%	2891 (2825,2957) 19%	3336 (3259,3411) 21%	3780 (3694,3866) 24%	4225 (4129,4321) 27%
RSV-bronchiolitis hospitalizations 2482(2336,2642) *					

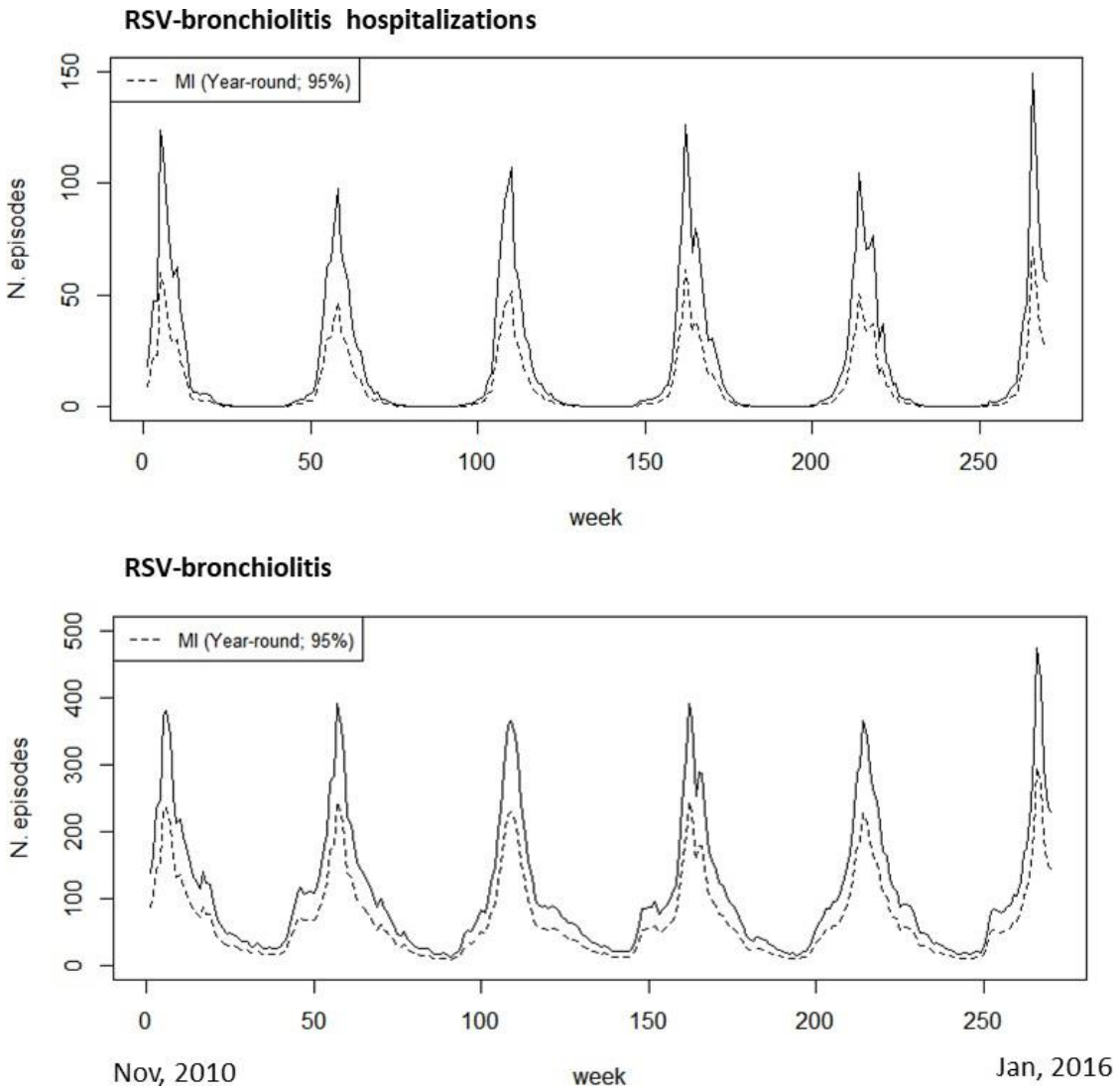
Seasonal	319 (299,341) 13%	377 (353,403) 15%	435 (407,464) 17%	493 (461,526) 20%	551 (516,588) 22%
Year-round	495 (466,528) 20%	586 (551,624) 24%	676 (636,719) 27%	766 (720,815) 31%	856 (805,911) 34%

*Annex 3.4_table_5: Impact of seasonal and year-round maternal immunization programs on RSV-bronchiolitis and RSV-bronchiolitis hospitalizations in sensitivity analysis 2. A) Provides the average number of bronchiolitis cases averted with the corresponding impact percentages, for the whole study period and B) by 100,000 infants per year. Different uptake MI levels (55% to 95%) were assessed. 95% CrI: 95% credible interval. * Number of predicted bronchiolitis without immunization program. ** The annual (October-September) average of number of averted bronchiolitis per 100,000 children was estimated excluding the incomplete periods 1) 2010-Nov, 2011-Sep and 2) 2015-Oct, 2016-Jan.*

Annex 3.5: Distribution of RSV-bronchiolitis by week under the different immunization scenarios

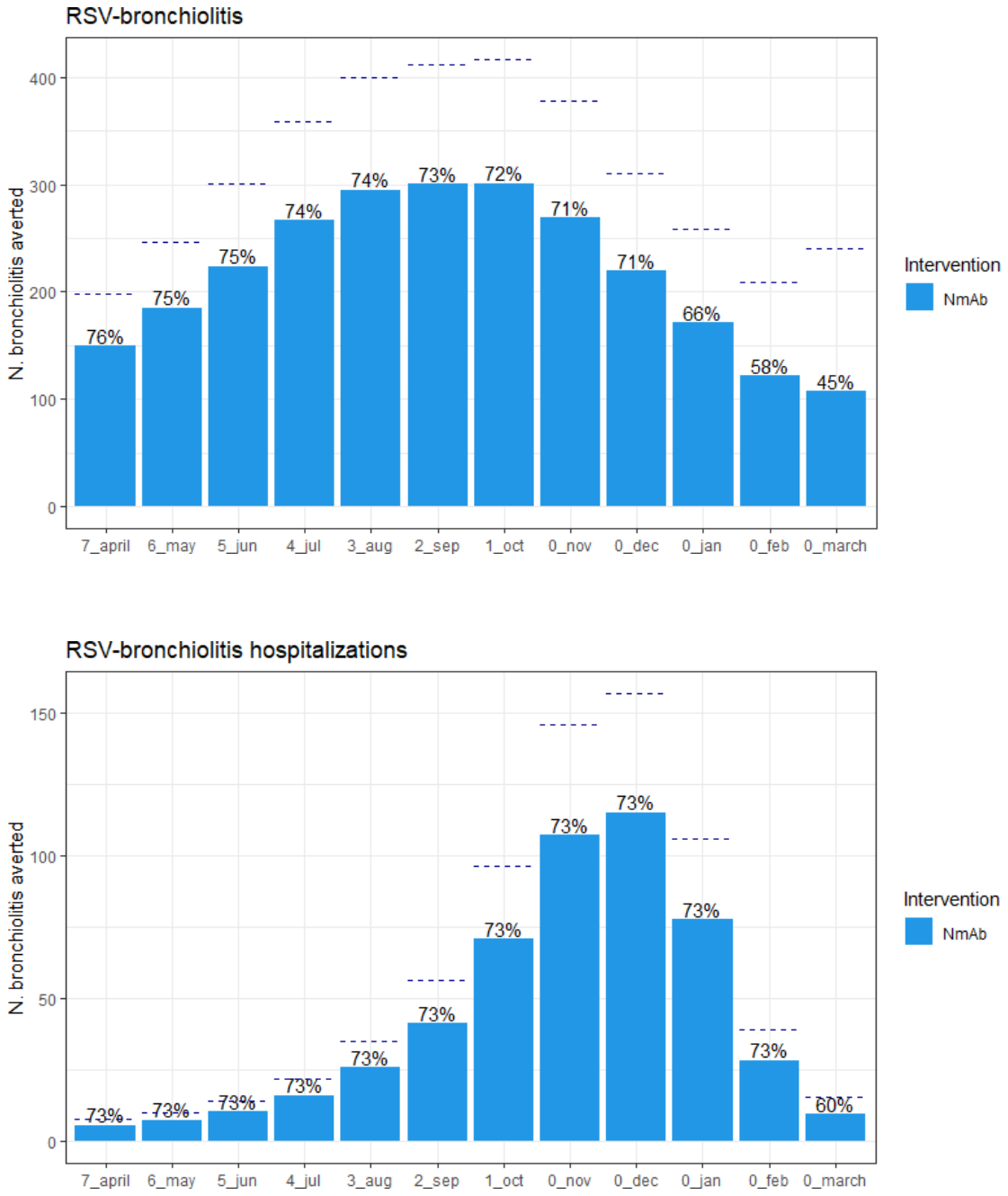


Annex 3.5a: weekly average counts of RSV-bronchiolitis and RSV-bronchiolitis hospitalizations predicted by the model without immunization strategies and under a seasonal with catch-up NmAb.



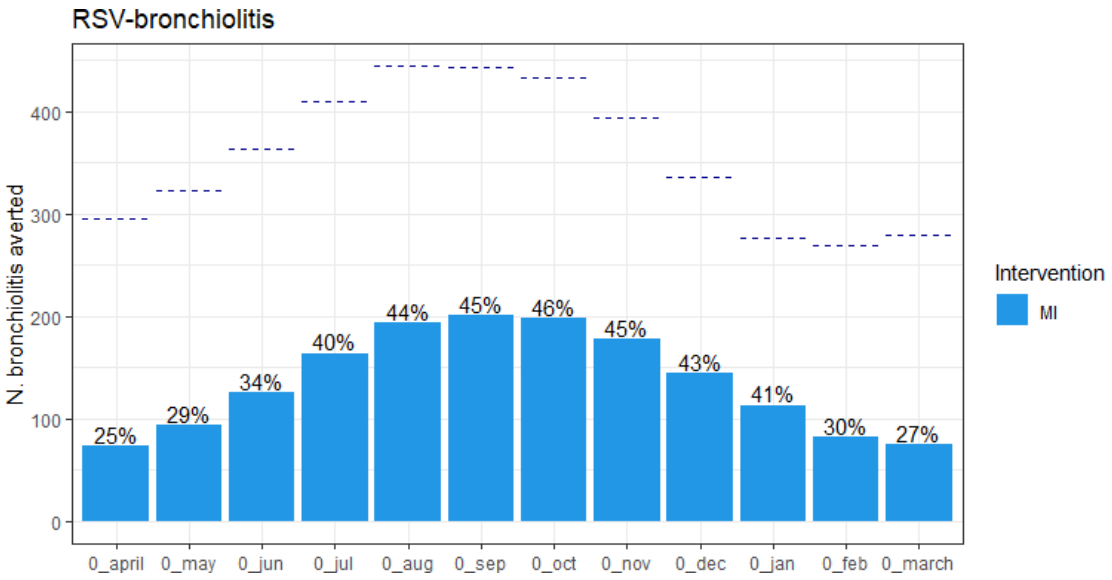
Annex 3.5b: weekly average counts of RSV-bronchiolitis and RSV-bronchiolitis hospitalizations predicted by the model without immunization strategies and under a year- round MI strategy.

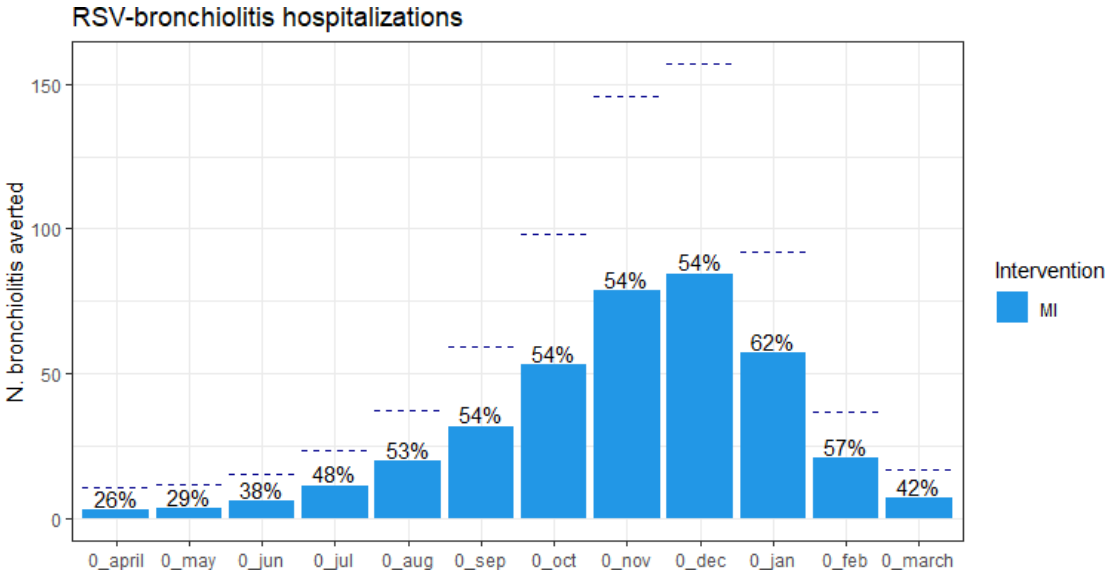
Annex 3.6: RSV-bronchiolitis averted by month of birth



Annex 3.6a: RSV-bronchiolitis and RSV-bronchiolitis hospitalizations averted, number and percentage, from November,2010- June, 2016 by month of birth under a seasonalwith catch-up NmAb. 95% coverage. X axis represents the age administration month of birth.

Dashed line: number of RSV-bronchiolitis predicted by the model without any of these interventions.





Annex 3.6b: RSV-bronchiolitis and RSV-bronchiolitis hospitalizations averted, number and percentage, from November,2010- June,2016 by month of birth under a year- round MI .95% coverage. X axis represents the age administration month of birth. Dashed line: number of RSV-bronchiolitis predicted by the model without any of these interventions.

Annex 4.1: Nimble code

```
SCluster_spacetime <- nimbleCode({

  for(i in 1:N_barrios){

    for (t in 1:6){

      y[i,t] ~ dbin(p[i,t],N[i,t])

      logit(p[i,t]) <- log(theta[i]) + gamma1*step(t - 7) +
                       (z_trend[i]==1)*(0*t + 0*(t - 7)*step(t - 7)) +
                       (z_trend[i]==2)*(beta21_dec*t*(1 - step(t - 7)) +
                       0*(t - 7)*step(t - 7)) +
                       (z_trend[i]==3)*(beta31_inc*t*(1 - step(t - 7)) +
                       0*(t - 7)*step(t - 7)) +
                (z_trend[i]==4)*(0*t + beta42_dec*(t - 7)*step(t - 7)) +
                       (z_trend[i]==5)*(beta51_dec*t*(1 - step(t - 7)) +
                        beta52_dec*(t - 7)*step(t - 7)) +
                       (z_trend[i]==6)*(beta61_inc*t*(1 - step(t - 7)) +
                        beta62_dec*(t - 7)*step(t - 7)) +
                (z_trend[i]==7)*(0*t + beta72_inc*(t - 7)*step(t - 7)) +
                       (z_trend[i]==8)*(beta81_dec*t*(1 - step(t - 7)) +
                        beta82_inc*(t - 7)*step(t - 7)) +
                       (z_trend[i]==9)*(beta91_inc*t*(1 - step(t - 7)) +
                        beta92_inc*(t - 7)*step(t - 7))

    }

    for (t in 7:9){

      y[i,t] ~ dbin(p[i,t],N[i,t])

      logit(p[i,t]) <- log(theta[i]) + gamma1*step(t - 7) +
                       (z_trend[i]==1)*(0*t + 0*(t - 7)*step(t - 7)) +
                       (z_trend[i]==2)*(beta21_dec*t*(1 - step(t - 7)) +
                        0*(t - 7)*step(t - 7)) +
                       (z_trend[i]==3)*(beta31_inc*t*(1 - step(t - 7)) +
                        0*(t - 7)*step(t - 7)) +
                (z_trend[i]==4)*(0*t + beta42_dec*(t - 7)*step(t - 7)) +
                       (z_trend[i]==5)*(beta51_dec*t*(1 - step(t - 7)) +
```

```
                    beta52_dec*(t - 7)*step(t - 7)) +
                   (z_trend[i]==6)*(beta61_inc*t*(1 - step(t - 7)) +
                    beta62_dec*(t - 7)*step(t - 7)) +
             (z_trend[i]==7)*(0*t + beta72_inc*(t - 7)*step(t - 7)) +
                   (z_trend[i]==8)*(beta81_dec*t*(1 - step(t - 7)) +
                    beta82_inc*(t - 7)*step(t - 7)) +
                   (z_trend[i]==9)*(beta91_inc*t*(1 - step(t - 7)) +
                    beta92_inc*(t - 7)*step(t - 7))

  }

  theta[i] <- inprod(z[i,1:k],eta[1:k])
  z[i,1:k] ~ dmulti(prlevels[1:k],1)
  z_trend[i] ~ dcat(prlevels_trend[1:k_trend])

}

gamma1 ~ dnorm(0,0.001)

# Increasing-decreasing trends

beta31_inc ~ T(dnorm(0,0.001), 0, 400)
beta61_inc ~ T(dnorm(0,0.001), 0, 400)
beta72_inc ~ T(dnorm(0,0.001), 0, 400)
beta82_inc ~ T(dnorm(0,0.001), 0, 400)
beta91_inc ~ T(dnorm(0,0.001), 0, 400)
beta92_inc ~ T(dnorm(0,0.001), 0, 400)

beta21_dec ~ T(dnorm(0,0.001), -400, 0)
beta42_dec ~ T(dnorm(0,0.001), -400, 0)
beta51_dec ~ T(dnorm(0,0.001), -400, 0)
beta52_dec ~ T(dnorm(0,0.001), -400, 0)
beta62_dec ~ T(dnorm(0,0.001), -400, 0)
beta81_dec ~ T(dnorm(0,0.001), -400, 0)

prlevels[1:k] ~ ddirch(alfa[1:k])
prlevels_trend[1:k_trend] ~ ddirch(alfa_trend[1:k_trend])

eta[1] <- increta[1]
for(j in 2:k){
  eta[j] <- eta[j-1] + increta[j]
```

```
  }

  for(j in 1:k){
    increta[j] ~ dgamma(1,1)
  }

  for(j in 1:k){
    prop[j] <- sum(z[1:N_barrios,j])
  }

})
```