ABSTRACT

DEPUY, VENITA MARIE. Modeling Transmission Dynamics of *Streptococcus pneumoniae* in the Presence of Vaccination: Model Development and Cost-Benefit Analysis. (Under the direction of Dr. Alun Lloyd.)

Streptococcus pneumoniae bacteria are a common cause of meningitis, bacteremia, and pneumonia, especially in young children. We developed an age-structured dynamic transmission model to explore the competition between different serotypes and the effect of vaccination on the incidence of colonization and pneumococcal disease among different age groups. Our model estimated that pneumococcal colonization would decrease from 12% to 7% within ten years following the introduction of the PCV7 vaccine in the United States, with the largest decreases occurring among young children. The number of invasive pneumococcal disease (IPD) cases is estimated to concurrently decrease from 28.7 to 16.3 cases per 100,000 subjects.

The relationship between antimicrobial resistance and inpatient mortality, length of stay, and cost among subjects with *S. pneumoniae* isolates from a sterile site was explored, using a sample of 578 patients from a single hospital system. Resistance was categorized as both dichotomous resistance (resistance to any antimicrobial agent) and incremental resistance (resistance to 0, 1, 2, or 3+ classes of antimicrobial agents). No relationship was found between inpatient mortality and antimicrobial resistance or isolate location. Length of stay was significantly related to both isolate location and antimicrobial resistance. Cost was significantly associated with race, isolate location, and antimicrobial resistance. Estimates of hospitalization costs and length of stay were provided by resistance level and significant baseline characteristics.

Average hospitalization costs were combined with model-based estimates of colonization and IPD to perform a cost-benefit analysis of the PCV7 vaccine in the United States. This analysis was focused on meningitis, bacteremia, and pneumonia costs during the first ten years following the introduction of the vaccine. Meningitis and bacteremia estimates were based on fractions of model-estimated IPD. Pneumonia estimates were calculated as fractions of model-estimated colonization. A hypothetical no-vaccine scenario, in which the

number of pneumococcal disease cases per 100,000 subjects remained constant at 1999 levels, was used to determine the number of cases of pneumococcal disease prevented by the vaccine. The total amount of prevented disease costs was calculated for 2000 – 2009, and the cost savings calculated as the prevented costs minus the costs of vaccination. We estimated that the vaccine resulted in a 63.7 billion dollar cost savings during the first ten years, or approximately $499 per dose administered. Additional analyses explored the additional predicted cost savings through 2025 and the use of a nationwide inpatient sample for average hospitalization costs. The vaccine was found to be cost-effective in both scenarios.

Modeling Transmission Dynamics of *Streptococcus pneumoniae* in the Presence of Vaccination: Model Development and Cost-Benefit Analysis

by
Venita Marie DePuy

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Biomathematics

Raleigh, North Carolina

2012

APPROVED BY:

Anastasios Tsiatis	Charles Smith
Alun Lloyd Chair of Advisory Committee	Jim Riviere

DEDICATION

For Jonathan.

BIOGRAPHY

The author was born and raised in McMinnville, Oregon. She especially enjoyed her science and math classes in high school, and attended the University of Alaska Fairbanks to further develop those interests. There, she received her baccalaureate degree in Statistics, with a minor in Biological Sciences.

Her interest in health care also grew during her time in Alaska, where she worked at Denali Center, a long-term care facility, while completing her degree. She then moved to North Carolina to pursue a Master of Statistics degree at North Carolina State University (NCSU). She was able to combine her health care interest and statistical background when she accepted a position as a statistician at the Duke Clinical Research Institute (DCRI) upon graduation.

One of the most interesting projects she worked on at DCRI involved investigating the relationship between antimicrobial resistance and hospitalization cost, length of stay, and mortality for different bacteria. This project, combined with the recent introduction of a vaccine for one of the bacteria (*S. pneumoniae*), interested her enough to want to investigate it further. She decided to pursue her doctorate part-time while continuing to work, and chose the Biomathematics program at NCSU.

While working on her doctorate, she has continued to work as a statistician, first at DCRI, then at a clinical research organization, and now enjoys working as a contract biostatistician, SAS programmer, and biomathematician through her own company, Bowden Analytics.

She also enjoys spending time with her family, and looks forward to having more time to travel.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABCs	Active Bacterial Core Surveillance
CAP	Community-acquired pneumonia
CBPD	U.S. Census Bureau Population Division
CCS	Clinical Classification Software
CDC	Centers for Disease Control and Prevention
CI	Confidence interval
CSF	Cerebrospinal fluid
GLM	Generalized linear model
HCUP	Healthcare Cost and Utilization Project
IPD	Invasive pneumococcal disease
LOS	Length of stay
MDR	Multi-drug resistant
NVSS	National Vital Statistics System
NVT	Non-vaccine-type (refers to PCV7 vaccine)
PCV13	13-serotype pneumococcal conjugate vaccine
PCV7	7-serotype pneumococcal conjugate vaccine
PRCC	Partial rank correlation coefficient
SE	Standard error
SIS	Susceptible-Colonized-Susceptible
SIRS	Susceptible-Colonized-Recovered-Susceptible
VT	Vaccine-type (refers to PCV7 vaccine)
WAIFW	'Who acquires infection from whom'
yo	Years old

LIST OF SYMBOLS

b_j, b_j*	Competition parameters (Zhang et al. model)
c_j	Competition within subject, by serotype group
c_j*	Competition parameters (Zhang et al. model)
d	Ratio of NVT:VT serotype transmission probabilities
d_{Xi}	Fraction of subjects who develop IPD and immediately die (Snedecor et al. model)
e	Effect of vaccine on colonization by VT serotypes
E_X	Equilibrium (E_0, E_1, E_2, E_{12})
f_{ij}	Fraction of colonized subjects who develop IPD, by serotype and age groups
g_i	Growth between age groups
h_j, h_j*	Competition parameters (Zhang et al. model)
I	Colonized (subscript)
k	Effect of vaccine on colonization by NVT serotypes
m_{ij}	Average weekly contact rate between i^{th} and j^{th} age groups
p_1	Transmission probability, VT, among <2 yo
p_2	Transmission probability, VT, among 2-15 yo and between <2 yo and 2-15 yo
p_3	Transmission probability, VT, among 16+ yo and btwn. 16+ and other age groups
r	Effect of vaccine on VT IPD
R	Recovered and temporarily immune (subscript)
R_{0j}	Basic reproductive number
S	Susceptible (subscript)
v	Vaccination fraction or vaccination rate
$v_j(t)$	Vaccination rates (Snedecor et al. model)
V	Vaccinated (subscript)
β^a	Assortative mixing parameter (Melegaro et al. model)
β_{ij}	Transmission parameter
β^p	Proportionate mixing parameter (Melegaro et al. model)
γ_{ji}^{-1}	Average duration of colonization, by serotype group and age group
δ_{xy}	Indicator variable, 1 when $x = y$ and 0 when $x \neq y$
ε, ε_i	Assortative/proportionate mixing weighting factor (Melegaro et al., Snedecor et al. models)
ζ_{Xi}	Fraction of col. subjects who develop IPD; X indicates disease type (Snedecor et al. model)

λ_{ji}, λ_j	Force of infection
μ_i	Death rate of i^{th} age group
μ_N	Birth rate
ρ_j^{-1}	Average duration of temporary immunity, by serotype group
ω^{-1}	Average duration of vaccine effect (Temime et al., Melegaro et al. models)

Note: i = 1, 2, 3 indicates age groups (typically <2 yo, 2-15 yo, 16+ yo); j = 1, 2 indicates serotype group (1=VT, 2=NVT)

1. INTRODUCTION

Streptococcus pneumoniae is a leading cause of pneumonia, bacteremia, and meningitis, and is a considerable public health concern worldwide. The World Health Organization has estimated that, in HIV-negative children less than 5 years of age, *S. pneumoniae* causes 14.5 million cases of serious illness and 735,000 deaths annually [1].

The complicated natural history of this bacterium makes it a challenging system to characterize. It is most commonly transmitted by asymptomatically colonized healthy individuals, making it difficult to assess transmission. Acquisition of the bacteria first leads to asymptomatic colonization, after which bacteria may spread throughout the body to cause infection in the lungs, blood, cerebrospinal fluid, or other areas. When the body's immune system reacts to the bacterial infection, such as the development of fluid in the lungs, it is considered pneumococcal disease.

The incidence of both colonization and pneumococcal disease vary by age group. The average number of contacts between people, and the subsequent chance of bacterial transmission, also varies between age groups. Therefore, a detailed model is necessary to incorporate the complexities of this system.

The large pneumococcal disease burden, and associated economic impact, has led to interest in developing *S. pneumoniae* vaccines to reduce the effect of the disease. This has subsequently led to interest in understanding the impact of the vaccine, including its cost-effectiveness.

Natural History of Pneumococcal Disease

S. pneumoniae is one of many bacterial flora normally found in the upper respiratory tract, especially among children. When the bacteria adhere to the epithelial lining of the respiratory tract, it is considered *colonization* [2]. Colonized subjects have no apparent symptoms, so must be identified by the use of a nasopharyngeal swab. These

asymptomatically colonized subjects serve as the primary source of transmission to other hosts, with bacteria typically spread through exposure to nasal aspirate, such as that from sneezing or coughing [3]. Exposure does not always lead to colonization [4]. Children have higher levels of colonization [2; 5-8], and longer average durations of colonization [9; 10] than adults. At least 90% of young children are colonized within the first two years of life [11-13].

If the bacteria penetrate the mucosal barrier and enter the tissues, it is considered pneumococcal *infection* [3]. This leads to pneumococcal disease, such as pneumonia, bacteremia, or meningitis. Different authors characterize pneumococcal disease in different ways. Pneumococcal disease present in a normally sterile site is considered invasive pneumococcal disease (IPD). The Centers for Disease Control and Prevention (CDC) does not consider pneumonia to be IPD, unless *S. pneumoniae* is also present in the blood (invasive pneumococcal pneumonia) [14]. An estimated 10 – 25% of pneumonia becomes bacteremic (i.e., spread to the bloodstream) [15], with the remaining disease categorized as non-invasive pneumonia. Other authors either consider pneumonia to be an invasive disease [2; 16] or consider pneumonia to be an infection of a normally sterile site [3]. In this research, IPD refers to invasive disease using the CDC definition, and pneumococcal disease refers to all illnesses caused by *S. pneumoniae*.

Since colonization is asymptomatic, and resource utilization and fatalities occur due to pneumococcal disease, the majority of research and data collection focuses on pneumococcal disease. The most consistent source of data regarding pneumococcal disease is the annual report of IPD incidence collected by the CDC Active Bacterial Core Surveillance (ABCs) [17]. The CDC ABCs Emerging Infections Program Network collects information from microbiology laboratories serving acute care hospitals in the surveillance area, which includes over 29 million people across ten states. Due to the collection method, some subjects in the surveillance area who have IPD may not be reported, such as subjects who are

not treated by an acute care hospital, hospitalized subjects whose isolates are not sent to a laboratory, or hospitalized subjects who do not have an isolate collected from the infected area.

The CDC annual *S. pneumoniae* surveillance report summarizes both the total IPD incidence and the rate of IPD (i.e., number of cases per 100,000 subjects in the surveillance area) by race, age group, infection type, and antimicrobial resistance. The rate of reported IPD, from 1999 to 2009, is shown by age group in Figure 1.1. The three age groups presented in this figure were consolidated from the eight age groups reported by the CDC ABCs [17].

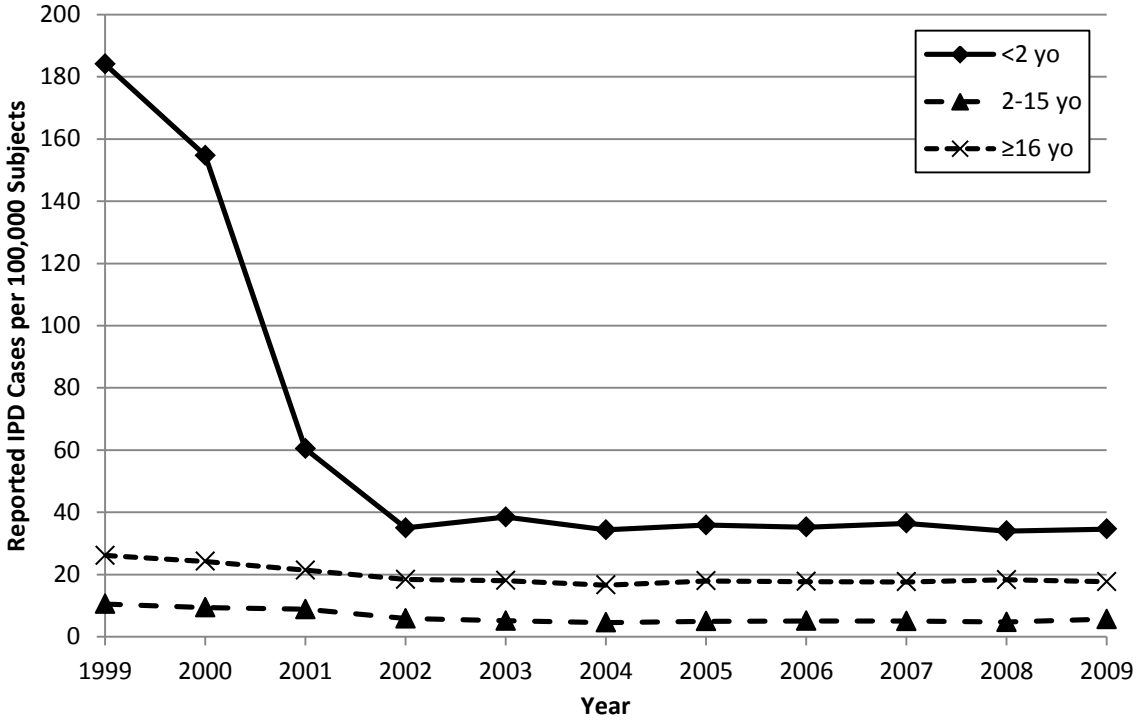


Figure 1.1
Annual Reported IPD Incidence by Age Group in the United States

Non-invasive pneumococcal disease, particularly pneumonia, is very prevalent in the United States. Pneumococcal pneumonia is poorly characterized since a determination of the cause of illness is not necessary for a diagnosis of pneumonia. Therefore, pneumococcal

pneumonia is often diagnosed as pneumonia due to an unspecified organism, making it difficult to accurately characterize its incidence. Most research to date has focused on all-cause community-acquired pneumonia (CAP), with specific information on the incidence, costs, and mortality associated with pneumococcal pneumonia generally lacking [18].

The CDC has estimated that *S. pneumoniae* causes ten times as many cases of pneumonia as bacteremia, and more than 150 times as many cases as meningitis [19]. Therefore, it is important to consider pneumonia, both invasive and non-invasive, when evaluating the impact of pneumococcal disease. Although disease rates have been reduced after the introduction of pneumococcal vaccines in recent decades, pneumococcal disease is still a major concern.

Vaccine Development

Due to the large number of people infected by *S. pneumoniae* annually, this bacterium has been a focus of vaccine development. Vaccines target certain serotypes of *S. pneumoniae*, which are distinguished from each other by the composition of the external polysaccharide capsule [20]. More than 90 different serotypes have been identified to date. These serotypes vary widely in geographical distribution, frequency of colonization, and ability to cause pneumococcal disease.

A 23-serotype polysaccharide vaccine for *S. pneumoniae* has been available for approximately 20 years. It is recommended for adults 65 years of age and older, as well as being recommended for people who have certain immunocompromising conditions and are at least 2 years of age. It has not been shown to produce an adequate immune response in children less than 2 years of age.

Due to the frequency and severity of IPD in young children, pediatric pneumococcal conjugate vaccines have been developed in recent years. The first pneumococcal conjugate vaccine (PCV7) was licensed for use in children in the United States in February 2000 [21]. This vaccine targets seven serotypes: 4, 6B, 9V, 14, 18C, 19F, and 23F. These serotypes

were selected for inclusion in the PCV7 vaccine because they have historically caused the majority of IPD in the United States. During 1998 and 1999, these seven serotypes caused 83% of IPD among children less than 2 years of age, and more than 50% of IPD among adults at least 20 years of age [22].

The introduction of the PCV7 vaccine led to a very large decrease in the total IPD among young children, and smaller declines among older age groups, as shown in Figure 1.1. It has also been shown to be effective against radiograph-confirmed pneumonia (much of which is not considered IPD) [23]. The decrease in total IPD has been driven by the reductions in IPD caused by the serotypes included in the vaccine [22; 24-30]. Similar reductions have also been observed in colonization by those serotypes [31-39]. The majority of studies have also observed a slight increase in colonization [31; 33-36; 38-40] and IPD [22; 24; 26-28] caused by the non-vaccine serotypes, a process called *serotype replacement*.

An expanded pneumococcal conjugate vaccine, PCV13, was licensed for use in children in the United States in February 2010 [41]. This vaccine includes serotypes 1, 3, 5, 6A, 7F, and 19A in addition to the seven serotypes in the PCV7 vaccine. Due to its recent introduction, very limited information is available on the impact of this vaccine. Therefore, our research focused on the PCV7 vaccine and its impact.

Modeling

Due to the spread of bacteria by healthy individuals, it is necessary to include colonization in the characterization of the *S. pneumoniae* epidemiological system. Certain parameters of this system, such as transmission probabilities and fractions of colonized subjects developing IPD, are typically unable to be directly determined using observational data. The indirect effects of vaccination on a population (i.e., herd immunity) are also difficult to quantify. These characteristics may, however, be estimated through modeling techniques.

Early mathematical models described a framework for pneumococcal colonization [42; 43]. These models used two-serotype systems to track the colonization and vaccination status of

subjects in the population. Direct competition between serotypes was included in these transmission models. In other words, if a subject is already colonized with one serotype, in the presence of direct competition that subject is less likely to become dually colonized with a second serotype than a completely susceptible subject would be. In order for serotype replacement to occur, as has been observed since the introduction of the vaccine, competition between serotypes must be present.

Temime et al. expanded that framework to model pneumococcal colonization in France, both before and after the introduction of the vaccine [44]. That model also included further refinements, such as accounting for patterns of interaction between subjects in different age groups (i.e., school-age children spend more time in contact with each other). Characterization of age structure can be complicated, due to the increased numbers of parameter values needed [45]. These authors predicted that, due to competition between serotypes and the resulting serotype replacement of vaccine-type serotypes by non-vaccine-type serotypes after vaccination, the vaccine would cause a very slight drop in total colonization levels, but no large decreases.

More recent epidemiological models have expanded the colonization framework to develop systems that include both colonization and IPD [46-48]. These models vary significantly in their characterization of vaccine effect, forces of infection, and relationships between different age groups.

We developed an age-structured dynamic transmission model, incorporating findings from recent literature, to estimate colonization and IPD caused by vaccine-type and non-vaccine-type serotypes in the United States, following the PCV7 introduction. This model was calibrated to IPD incidence in 1999 – 2009, as reported by the CDC [17], and was used to estimate colonization and pneumococcal disease incidence in the United States. This model also fitted unknown parameter values, such as transmission probabilities and fractions of colonized subjects developing IPD, which are otherwise difficult to estimate. We then

explored the economic impact of pneumococcal disease, and later incorporated those cost estimates into model-based estimates of pneumococcal disease to further evaluate the effects of vaccination.

Economic Impact and Antimicrobial Resistance

It is important to estimate the incidence of pneumococcal disease because of the significant economic burden caused by *S. pneumoniae*. This bacterium is a leading cause of CAP [49], as well as being the leading cause of bacterial meningitis [50]. Patients with pneumonia have also been shown to present a significant financial burden to employers [51], with the majority of costs attributed to hospitalized patients.

Several studies have examined certain types of pneumococcal disease, but often included all-cause CAP, or were limited to specific age groups, infection sites, or antimicrobial resistance types. Cost analyses of pneumococcal disease are further complicated by changes over time, as different treatments and different medical practices emerge; there are also increasing trends in antimicrobial resistance and hospitalization costs.

Antimicrobial resistance occurs when a bacterium exhibits decreased susceptibility to a specific group of antimicrobial agents. The first known cases of penicillin-resistant *S. pneumoniae* cultures were identified in the 1960s. Antimicrobial resistance rates have typically been found to be increasing over time [52-56], and vary by geographic area [57]. For instance, in Cleveland, Ohio, the first penicillin-resistant strain was isolated in 1980 and the first macrolide-resistant strain in 1984; by 2004, over 50% of isolates were penicillin-resistant and/or macrolide-resistant [55].

In general, infections caused by antimicrobially resistant bacteria are believed to cause longer hospitalization times and higher costs when compared with infections by susceptible strains of bacteria [58]. Resistance can affect patient outcomes by limiting available treatment options, causing a delay in the administration of appropriate therapy, and enhancing virulence [59]. Broad-spectrum antimicrobials, such as fluoroquinolones or third-generation

cephalosporins, are now often required for treatment of infections; these agents are typically more expensive, have a greater impact on protective microflora, and can be more toxic or less effective [59]. Furthermore, a patient with a history of resistant bacterial infections may be treated with a very strong antimicrobial agent, such as vancomycin, when a more narrow-spectrum agent might have worked for that particular infection. This increased use can also contribute to the growth of resistance against these newer, broad-spectrum antimicrobials.

The cost of medical care for specific conditions is a key element of pharmacoeconomics. With so many variables contributing to the medical cost when patients are admitted to the hospital – such as antimicrobial resistance and site of infection – it is often hard to evaluate the expected direct medical cost, which is of vital importance to researchers, insurance providers, vaccine manufacturers, and other stakeholders. Accurate estimates of inpatient cost, mortality, and length of stay (LOS) are vital to the development of economic models to assess the effectiveness of new treatments [60].

There is very little information available about the direct effects of antimicrobial resistance on the costs of pneumococcal disease [18]. While some researchers have presented summaries of cost, LOS, and/or mortality in patients with *S. pneumoniae* infections or with CAP in general, these figures typically present average values for that sample, leading to estimates that are dependent on the demographic and disease characteristics of the patients in that sample, without controlling for covariates. In addition to further exploring the relationships between increasing levels of antimicrobial resistance and cost, LOS, and mortality, we provide estimates of cost and LOS based on the demographic and disease characteristics shown to be significantly related to each outcome. This allows us to provide more accurate estimates of the economic impact of the hospitalization of a patient with certain baseline characteristics.

Cost Effectiveness of Vaccination

Another key area of economic interest is the cost-effectiveness of the pneumococcal conjugate vaccines. The PCV7 and PCV13 conjugate vaccines are some of the most expensive vaccines introduced in the United States, with private pay costs from the CDC price list of $83.88 (PCV7) and $108.75 (PCV13) in 2010 [61]. For comparison, the 38 pediatric vaccines listed on the current CDC price list have an average private pay cost of less than $40.

The first cost-effectiveness study of PCV7 for the United States, published in 2000, concluded that the PCV7 vaccine would prevent more than 65,000 cases of meningitis, bacteremia, and pneumonia, and would prevent 116 deaths due to pneumococcal infection in a hypothetical birth cohort of 3.8 million people, but it would not be cost effective at the list price [62]. Another cost-effectiveness analysis compared two hypothetical birth cohorts in the United Kingdom – one vaccinated, one unvaccinated – and concluded that the vaccine cost would need to be reduced by approximately half in order to make the cost per quality-adjusted life-year fall within acceptable limits [63]. These single-cohort analyses did not incorporate the indirect effects of vaccination. For example, a vaccine that reduces colonization would then reduce overall transmission in the population, and reduce the incidence of colonization among unvaccinated individuals.

A more recent analysis, incorporating data from the first seven years of vaccination in the United States, estimated that the characterization of indirect effects of vaccination reduced the estimated cost per life-year saved by 95% [64]. Those researchers concluded that PCV7, with an estimated cost of $10,400 per life-year saved, has likely been a cost-effective vaccine when compared to the cost-effectiveness of other medical interventions currently in use.

We used our dynamic transmission model to estimate the incidence of meningitis, bacteremia, and pneumonia since the introduction of the PCV7 vaccine, and to estimate the

amount of disease prevented. We then calculated the cost savings due to prevented disease, and performed a cost-benefit analysis.

Organization of Dissertation

This research is organized as follows. Chapter 2 reviews previously published dynamic transmission models for *S. pneumoniae* [42-44; 46; 47]. Chapter 3 describes the development of an age-structured dynamic transmission model that uses parameter values from available literature and is calibrated to fit reported IPD incidence. Chapter 4 investigates the relationship between antimicrobial resistance in *S. pneumoniae* and the average duration of inpatient hospitalization, cost, and mortality, and provides estimates based on demographic and disease characteristics. Chapter 5 uses model-based estimates of colonization and IPD, in conjunction with average hospitalization costs, to evaluate the cost-effectiveness of the PCV7 vaccine. Chapter 6 summarizes our findings and suggests future research.

2. RECENT DYNAMIC TRANSMISSION MODELS FOR *S. PNEUMONIAE*

Several dynamic transmission models for *Streptococcus pneumoniae* have been previously developed. Lipsitch [42] introduced a two-serotype colonization model in which subjects were susceptible to (S) or colonized by (I) each of the two serotypes. In addition, subjects could be vaccinated (V) against the first serotype. This SIS model provided a framework for analyses of equilibria and model dynamics. Zhang, Auranen, and Eichner [43] expanded that framework to a two-serotype SIRS colonization model with the incorporation of a temporarily immune (R) compartment for each serotype.

Temime, Guillemot, and Boëlle [44] expanded Lipsitch's SIS colonization model to include age structure and the effect of antibiotic treatment. Serotypes were categorized into two groups (those included in the newly introduced vaccine, and those not included), and real-life parameter values were used. The model was then used to estimate colonization following the introduction of the vaccine. Melegaro et al. [46] and Snedecor et al. [47] later developed age-structured models incorporating both colonization and invasive pneumococcal disease (IPD). These models also used observed values to calculate parameter estimates, and predicted IPD following the introduction of the vaccine.

Each of these models is reviewed below. The notation for each model was changed to be consistent between models. In all cases, the models tracked the fractions of the population in each compartment.

Two-serotype SIS Model with Vaccination (Direct Competition)

Lipsitch [42] published the first of the applicable models to explore the competition between two serotypes of *S. pneumoniae* and the effect of a serotype-specific vaccine on the persistence of colonization in the population.

The introduction of the *Haemophilus influenzae* (Hib) vaccine in 1985, followed by the introduction of the pediatric pneumococcal conjugate vaccine (PCV7) in 2000, led to speculation on the effects of competition between different serotypes of the same organism in the presence of a vaccine. For example, if competition exists between serotypes, and the prevalence of one serotype is reduced due to vaccination, the prevalence of the other serotype may increase. Several studies, both before and after the Lipsitch paper was published, have shown some degree of increase in colonization levels of serotypes not targeted by the vaccine [34; 38; 65-67], confirming the assumption of competition between serotypes.

Model Structure

Subjects in this two-serotype SIS model were susceptible to or colonized with one or both serotypes. Subjects also could have been vaccinated against one serotype, and been either susceptible or colonized with the other serotype. For consistency between descriptions of different models, we have referred to the first serotype as the vaccine-type (VT) serotype and the second serotype as the non-vaccine-type (NVT) serotype. We described each subject's colonization and vaccination status using a two-letter subscript, where the first subscript denotes the VT serotype and the second subscript denotes the NVT serotype. The first subscript was also used to indicate vaccination, which provides complete protection against the VT serotype. The total population was then divided into four compartments for unvaccinated subjects (N_{SS}, N_{SI}, N_{IS}, and N_{II}) and two compartments for vaccinated subjects (N_{VS} and N_{VI}). N_{SI} consisted of individuals who were susceptible to the VT serotype and colonized with the NVT serotype; this was distinct from N_{IS}, in which subjects were colonized with VT and susceptible to NVT serotypes. Time was measured relative to the average human lifespan; hence per-capita birth and death rates were taken to be 1. Vaccination was assumed to occur at birth, and provided permanent protection against colonization with the VT serotype, but may or may not have been effective against the NVT serotype. A fraction v entered a vaccinated compartment, N_{VS}, at birth, where they were susceptible to the NVT serotype. The remaining fraction 1- v of newborns entered compartment N_{SS}, where those individuals were susceptible to both serotypes.

Vaccinated individuals could become colonized with the NVT serotype (N_{VI}) and later become resusceptible to that serotype (N_{VS}). Unvaccinated individuals could become colonized with VT and NVT serotypes at rates λ_1 and λ_2, respectively, where $\lambda_1 = \beta_1(N_{IS} + N_{II})$ and $\lambda_2 = \beta_2(N_{SI} + N_{II} + N_{VI})$ and each N_{XX} represented a fraction of the population. Subjects were taken to recover from colonization with a given serotype at a constant rate γ, independent of vaccination.

When an individual was already colonized with one serotype, the reduction in the rate of colonization by the other serotype was characterized by a competition parameter c_i. Possible values of the competition parameters ranged from 0 to 1, where 1 indicated no reduction in colonization and 0 indicated that dual colonization cannot occur. The parameter k, whose possible values also ranged from 0 to 1, was used to indicate the reduction in NVT colonization when the subject had been vaccinated against the VT serotype. When $k = 0$, vaccination protected completely against both serotypes; when $k = 1$, colonization with the NVT serotype was unchanged by the presence of vaccination. Values of $k < 1$ indicated that the vaccine was at least partially effective against the NVT serotype (a *bivalent* vaccine).

Movement between compartments in this model is summarized by the equations shown in {2.1}, and displayed graphically in Figure 2.1:

$$
\begin{aligned}
dN_{SS}/dt &= (1-v) - \lambda_1 N_{SS} - \lambda_2 N_{SS} + \gamma N_{SI} + \gamma N_{IS} - N_{SS} \\
dN_{SI}/dt &= \lambda_2 N_{SS} + \gamma N_{II} - c_1\lambda_1 N_{SI} - \gamma N_{SI} - N_{SI} \\
dN_{IS}/dt &= \lambda_1 N_{SS} + \gamma N_{II} - c_2\lambda_2 N_{IS} - \gamma N_{IS} - N_{IS} \\
dN_{II}/dt &= c_1\lambda_1 N_{SI} + c_2\lambda_2 N_{IS} - 2\gamma N_{II} - N_{II} \\
dN_{VS}/dt &= v - k\lambda_2 N_{VS} + \gamma N_{VI} - N_{VS} \\
dN_{VI}/dt &= k\lambda_2 N_{VS} - \gamma N_{VI} - N_{VI}
\end{aligned}
\qquad \{2.1\}
$$

where $\lambda_1 = \beta_1(N_{IS} + N_{II})$ and $\lambda_2 = \beta_2(N_{SI} + N_{II} + N_{VI})$.

Equilibria

The basic reproductive number of each serotype, R_{0i}, was defined as the average number of secondary colonizations caused by a colonized individual in an otherwise entirely susceptible

population. In the absence of a vaccine, the prevalence of a serotype will decrease over time if $R_{0i} < 1$, and that serotype will persist in the population if $R_{0i} > 1$. R_{v1} denotes the number of secondary infections in a vaccinated population. If a fraction v of subjects were vaccinated against a specific (VT) serotype, the average number of secondary colonizations (R_{01}) would need to be higher than 1 for that serotype to persist in the population: $R_{v1} = R_{01}(1-v) > 1$.

Analysis of Lipsitch's model found four possible outcomes:

i. E_0: At equilibrium, $N_{SS} + N_{VS} = 1$, where $N_{SS} = 1-v$ and $N_{VS} = v$. This equilibrium was stable if $R_{v1} < 1$ and $R_{02} < 1/(v(k-1)+1)$. In other words, the average number of secondary colonizations produced by subjects colonized with either serotype, after accounting for vaccination, must be less than one.

ii. E_1: At equilibrium, $N_{SS} + N_{VS} + N_{IS} = 1$, where $N_{VS} = v$, $N_{SS} = 1/R_{01}$, and $N_{IS} = 1-v-1/R_{01}$. This equilibrium was present and biologically feasible when $R_{v1} > 1$ and $R_{02} < 1/(v(k-1)+1)$, and in addition was stable when $R_{01}/R_{02} > 1-c_2+R_{01}(v+c_2(1-v))$.

iii. E_2: At equilibrium, $N_{SS} + N_{VS} + N_{SI} + N_{VI} = 1$. This slightly more complex outcome is of a form intuitive from the E_1 results. A fraction v of the population was in the vaccinated subgroup, with $N_{VS} = v/R_{02}$ and $N_{VI} = v(1-1/R_{02})$. The same proportions held true in the unvaccinated subgroup of size $1-v$, with $N_{SS} = (1-v)/R_{02}$ and $N_{SI} = (1-v)(1-1/R_{02})$. This equilibrium was present and biologically feasible when $R_{02} > 1/(v(k-1)+1)$ and $R_{v1} < 1$. The equilibrium was also stable when vaccination was high enough to prevent coexistence, which was determined by calculating the necessary vaccination level to ensure that all eigenvalues for the Jacobian matrix of partial derivatives of the differential equations, evaluated at N_{SS}, N_{VS}, N_{SI}, and N_{VI} as described above, were negative. These equations are shown below in {2.2} and {2.3}.

iv. E_{12}: the coexistence of the two serotypes at equilibrium was possible if both serotypes had $R_{0i} > 1$ (or $R_{v1} > 1$ if vaccine was present) and each could invade the population when the other serotype was present at equilibrium. If no vaccine was present, or the vaccine

was VT specific (k = 1), this equilibrium was stable if $R_{v1} > 1$, $R_{02} > 1/(v(k\text{-}1)+1)$, and both inequalities in {2.2} were true.

$$\frac{1}{(1-c_1+c_1R_{02})(1-v)} < \frac{R_{01}}{R_{02}} < 1-c_2+R_{01}(v+c_2(1-v)) \qquad \{2.2\}$$

In the presence of a bivalent vaccine ($0 \leq k < 1$ and $v > 0$), the two serotypes coexisted when:

$$\frac{R_{01}}{R_{02}} < 1-c_2+R_{01}(kv+c_2(1-v)) \qquad \{2.3\}$$

and vaccination was below a certain threshold, determined by setting the eigenvalues in the Jacobian matrix at this equilibrium equal to zero and solving the resulting quadratic for the vaccination fraction. It should be noted that equation {2.3} is a correction to the equation displayed in the Lipsitch paper (correspondence with the original author has verified this correction).

A summary of these outcomes is shown in Figure 2.2, where columns reflect the values of the competition parameters and rows represent the presence of vaccination and its effect on the NVT serotype. These figures depict the regions of R_{01}, R_{02} parameter space where various outcomes were observed. When competition between serotypes prevented dual colonization (c_i = 0) and the vaccination either did not affect the NVT serotype (k = 1) or was completely absent (v = 0), the only time that both serotypes could coexist at equilibrium was when $R_{01} = R_{02}$.

Outcomes

The Lipsitch paper explored the basic structure of a two-serotype SIS model with vaccination, and provided insight into the incidence of each serotype at equilibrium for different parameter

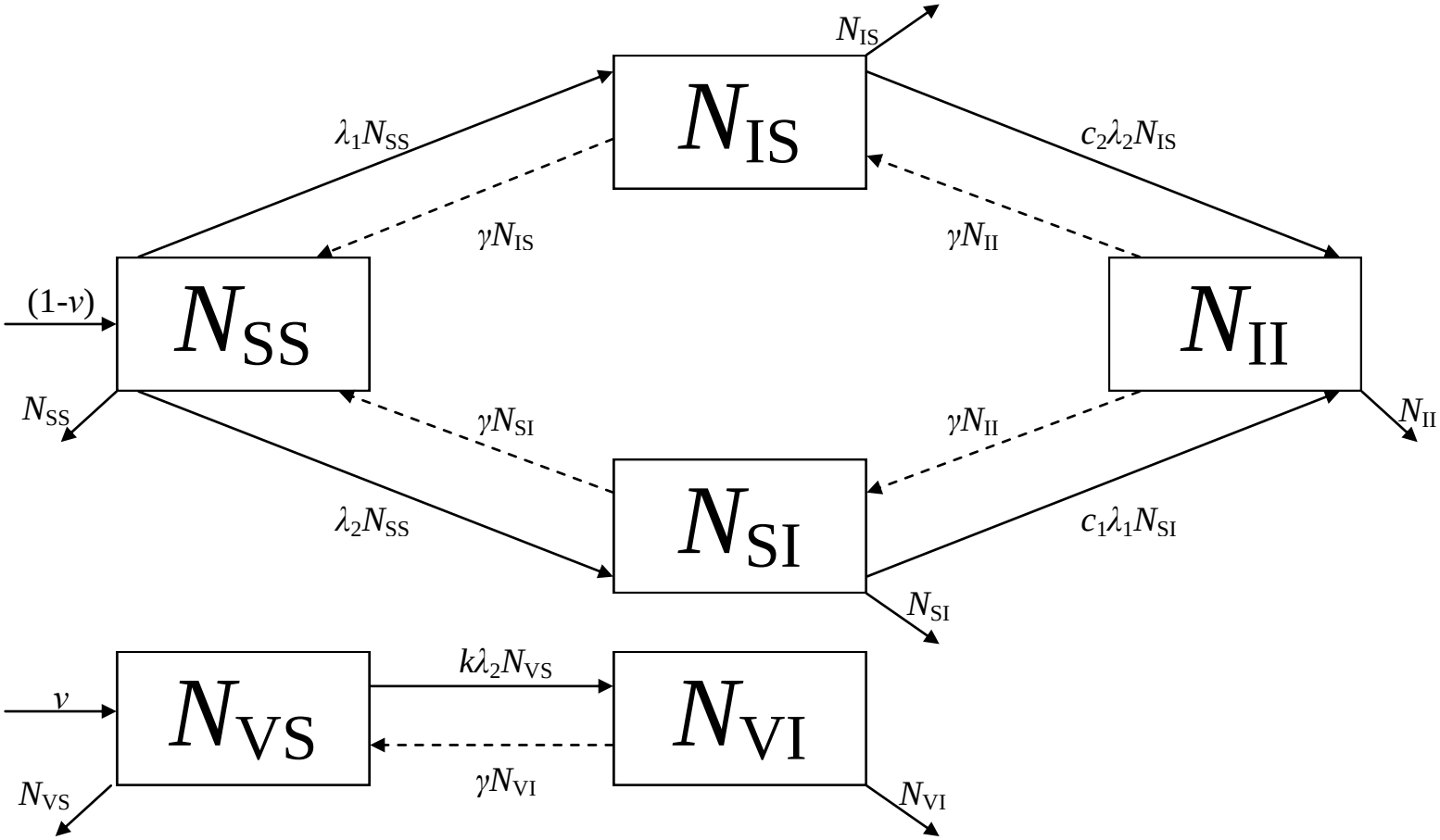
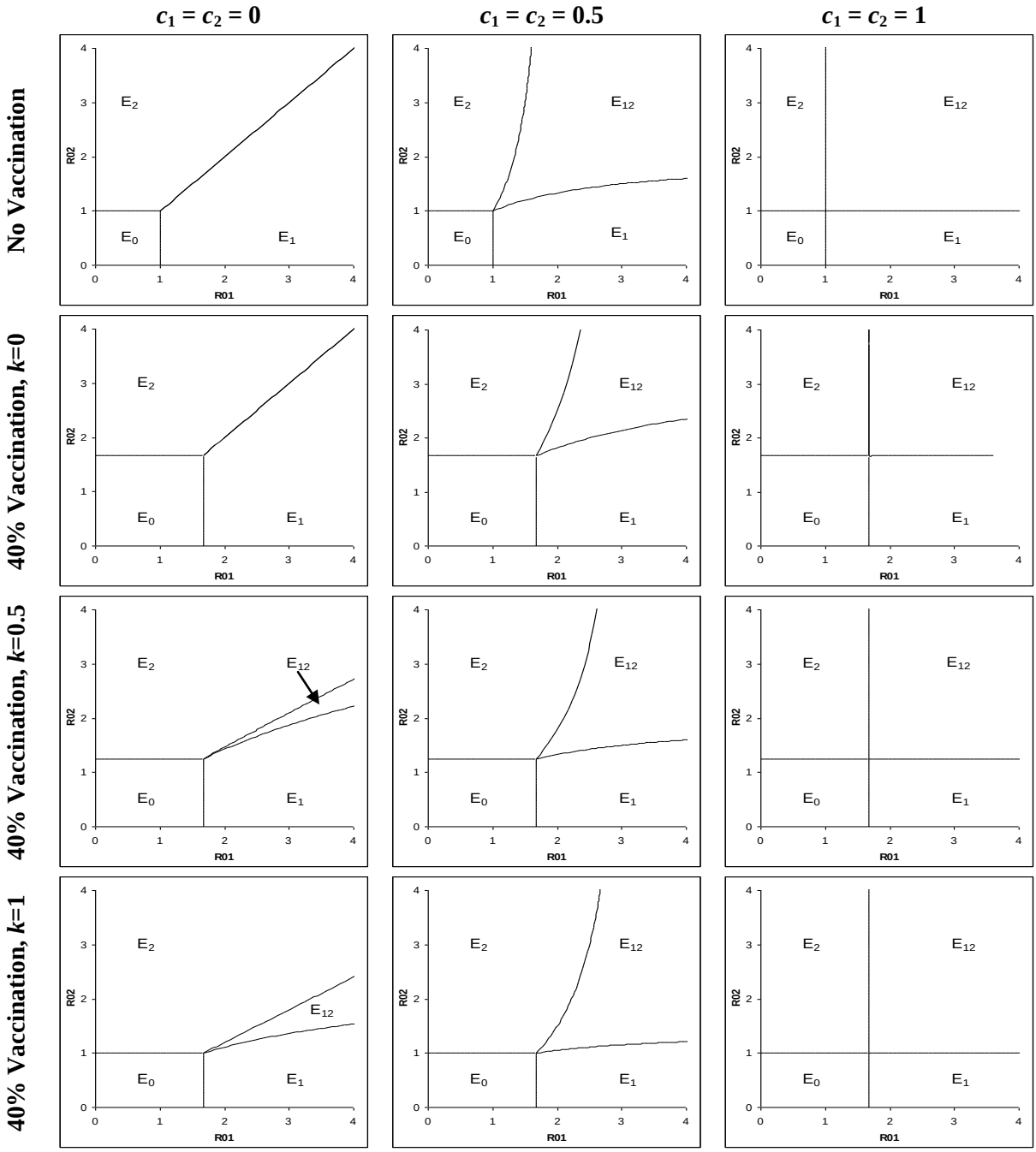


Figure 2.1
Two-Serotype SIS Model with Vaccination (Lipsitch)



Note: E_0 indicates neither serotype was present at stable equilibrium; E_1 and E_2 indicate the presence of the vaccine-type or non-vaccine-type serotype only at stable equilibrium, respectively; and E_{12} indicates the dual-serotype stable equilibrium.

Figure 2.2
Pattern of Outcomes: Two-Serotype SIS Model with Vaccination (Lipsitch)

values. When no dual colonization was permitted ($c_i = 0$), the dual-serotype equilibrium was only present when $R_{01} = R_{02}$. When a serotype-specific ($k = 1$) vaccine was present, less vaccination was necessary to cause eradication of that serotype when competition between serotypes was high.

This model was simple enough to allow an in-depth investigation of the relationships between equilibria and parameter values, and provided the basis for future models. Later models expanded this framework, with the addition of age structure, temporary immunity, or alternative characterization of vaccination, leading to models that could be better applied to real-world data.

Two-serotype SIRS Model with Vaccination (Direct or Indirect Competition)

Zhang et al. [43] investigated the impact of direct and indirect competition, using models very similar to that used by Lipsitch. They expanded the two-serotype SIS model to include a period of temporary immunity following colonization, and evaluated the influence of competition and vaccination on the coexistence of two pneumococcal serotypes.

The direct competition model was very similar to the Lipsitch model previously discussed, with the addition of a temporarily immune (R) compartment and additional potential effects of dual colonization (increased recovery rate and reduced infectiousness) included. Although the model did not require that birth and death rates be equal to each other, the authors simplified the model by taking them to be equal. The model was also simplified slightly by characterizing vaccination as having no effect on NVT serotype colonization, effectively setting $k = 1$.

The indirect competition model also shared these traits but characterized competition differently. While the direct competition model included terms for effects on one serotype when the subject was currently colonized with the other serotype, the indirect model included

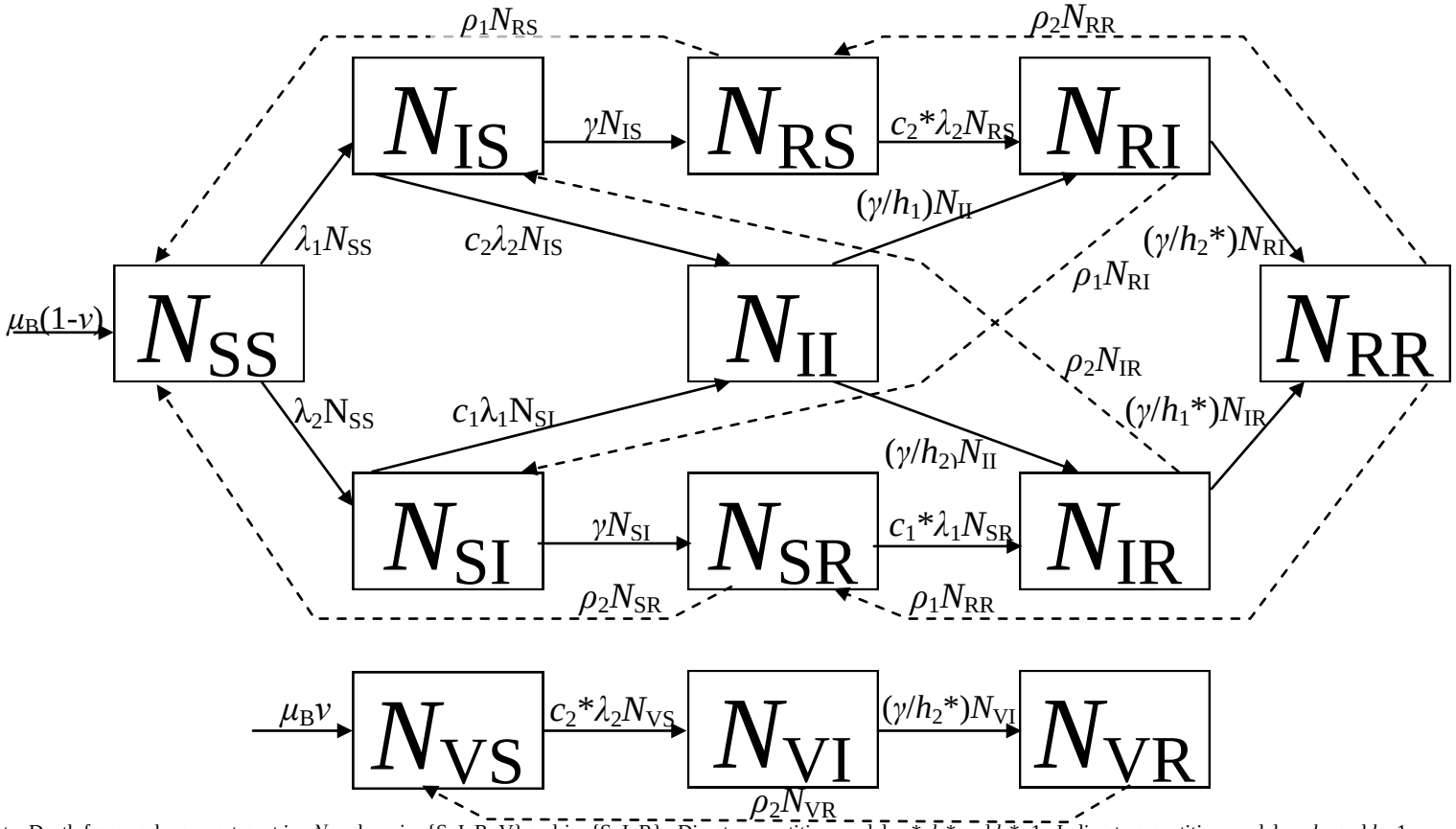
terms for effects when the subject was recovered from the other serotype, and presumably had antibodies in his system from the earlier colonization.

Model Structure

Both the direct and indirect competition models utilized 12-compartment SIRS models. Each subject was susceptible, colonized, or temporarily immune to each serotype, and subjects could be vaccinated against the VT serotype. Temporarily immune individuals became resusceptible at rates ρ_1 and ρ_2 for the VT and NVT serotypes, respectively.

The direct competition model included parameters indicating reduced colonization among subjects already colonized with one serotype (c_1 and c_2), as were present in the Lipsitch model. Additional competition terms (h_1 and h_2, b_1 and b_2), which resulted in potentially increased recovery rates and decreased infectiousness, respectively, were also present in the direct competition model. The indirect competition model incorporated slightly different competition terms (c_1* and c_2*, h_1* and h_2*, b_1* and b_2*), reflecting the effect of antibody-mediated competition. The decrease in NVT colonization due to vaccination, characterized as k in the Lipsitch model, was set equal to 1 (no reduction in colonization) in the direct competition model and set equal to c_2* in the indirect competition model. In other words, these two models treated vaccination either as not affecting colonization by the NVT serotype (direct competition model), or as affecting colonization by the NVT serotype by treating the immune response to vaccination as being identical to the immune response following VT colonization, through antibody-mediated competition (indirect competition model).

The equations shown in {2.4} describe both the direct and indirect models. For the direct competition model, parameters c_i*, h_i*, and b_i* were set equal to one. For the indirect competition model, parameters c_i, h_i, and b_i were set equal to one. Movement between compartments is shown graphically in Figure 2.3.



Note: Death from each compartment is μN_{ij}, where i = {S, I, R, V} and j = {S, I, R}. Direct competition model: $c_i{*}$, $b_i{*}$ and $h_i{*}$=1. Indirect competition model: c_i, b_i, and h_i=1.

Figure 2.3
Two-Serotype SIRS Model with Vaccination (Zhang et al.)

$$
\begin{aligned}
dN_{SS}/dt &= \mu_B(1-\nu) - \lambda_1 N_{SS} - \lambda_2 N_{SS} + \rho_1 N_{RS} + \rho_2 N_{SR} - \mu N_{SS} \\
dN_{SI}/dt &= \lambda_2 N_{SS} + \rho_1 N_{RI} - c_1\lambda_1 N_{SI} - \gamma N_{SI} - \mu N_{SI} \\
dN_{SR}/dt &= \gamma N_{SI} + \rho_1 N_{RR} - c_1 * \lambda_1 N_{SR} - \rho_2 N_{SR} - \mu N_{SR} \\
dN_{IS}/dt &= \lambda_1 N_{SS} + \rho_2 N_{IR} - c_2\lambda_2 N_{IS} - \gamma N_{IS} - \mu N_{IS} \\
dN_{II}/dt &= c_1\lambda_1 N_{SI} + c_2\lambda_2 N_{IS} - (\gamma/h_1) N_{II} - (\gamma/h_2) N_{II} - \mu N_{II} \\
dN_{IR}/dt &= c_1 * \lambda_1 N_{SR} + (\gamma/h_2) N_{II} - \rho_2 N_{IR} - (\gamma/h_1{}^*) N_{IR} - \mu N_{IR} \\
dN_{RS}/dt &= \gamma N_{IS} + \rho_2 N_{RR} - \rho_1 N_{RS} - c_2 * \lambda_2 N_{RS} - \mu N_{RS} \\
dN_{RI}/dt &= c_2 * \lambda_2 N_{RS} + (\gamma/h_1) N_{II} - \rho_1 N_{RI} - (\gamma/h_2{}^*) N_{RI} - \mu N_{RI} \\
dN_{RR}/dt &= (\gamma/h_1{}^*) N_{IR} + (\gamma/h_2{}^*) N_{RI} - \rho_1 N_{RR} - \rho_2 N_{RR} - \mu N_{RR} \\
dN_{VS}/dt &= \mu_B \nu + \rho_2 N_{VR} - c_2 * \lambda_2 N_{VS} - \mu N_{VS} \\
dN_{VI}/dt &= c_2 * \lambda_2 N_{VS} - (\gamma/h_2{}^*) N_{VI} - \mu N_{VI} \\
dN_{VR}/dt &= (\gamma/h_2{}^*) N_{VI} - \rho_2 N_{VR} - \mu N_{VR}
\end{aligned}
\qquad \{2.4\}
$$

where $\lambda_1 = \beta_1(N_{IS} + b_1 N_{II} + b_1{}^* N_{IR})$ and $\lambda_2 = \beta_2(N_{SI} + b_2 N_{II} + b_2{}^* N_{RI} + b_2{}^* N_{VI})$.

It is important to recall that the asterisk is part of the parameter notation, differentiating competition parameters between direct and indirect competition models, and does not indicate multiplication.

Equilibria

The possible outcomes for the direct and indirect models were very similar to those discussed in the Lipsitch model:

i. E_0: $N_{SS} = (\mu_B/\mu)(1-\nu)$ and $N_{VS} = (\mu_B/\mu)\nu$ at equilibrium.

ii. E_1: $N_{SS} + N_{VS} + N_{IS} + N_{RS} = (\mu_B/\mu)$ at equilibrium, where $N_{VS} = (\mu_B/\mu)\nu$ and $N_{SS} = 1/R_{01}$. The fraction of non-susceptible subjects colonized versus temporarily immune depended on the specific parameter values, with $N_{IS} = ((\rho_1 + \mu)/\gamma)\, N_{RS}$.

iii. E_2: $N_{SS} + N_{VS} + N_{SI} + N_{SR} + N_{VI} + N_{VR} = 1$ at equilibrium, where $N_{SS} = (1-\nu)/R_{02}$ and $N_{VS} = \nu/R_{02}$. The ratio of colonized to temporarily immune subjects, in both vaccinated and unvaccinated groups, was $(\rho_2 + \mu)/\gamma$. In the indirect competition

model, vaccination was parameterized as competition, so the equilibrium also depended on the competition terms c_2* and h_2*. In both cases, $N_{SS} + N_{SI} + N_{SR} = (\mu_B/\mu)(1-v)$ and $N_{VS} + N_{VI} + N_{VR} = (\mu_B/\mu)v$ at equilibrium.

iv. E_{12}: both the VT and NVT serotypes are present at equilibrium, when $R_{0i} > 1$ (or $R_{v1} > 1$ if vaccine was present). The presence of this equilibrium can be determined numerically, but the system cannot be explicitly solved algebraically.

Outcomes

The authors concluded that outcomes from each type of competition within model type (c_i, b_i, and h_i or c_i*, b_i*, and h_i*, depending on the model) were broadly similar. Longer average durations of immunity were associated with decreased levels of colonization, and a reduced impact of competition parameters on colonization levels. Direct competition had the greatest influence on colonization levels if the average duration of temporary immunity was short, and indirect competition had the largest impacts on colonization levels in the presence of long average durations of temporary immunity.

Although this model expanded the framework presented in the Lipsitch model, the intent was not to incorporate observed data, or to provide model-based estimates for real-world scenarios. The following model incorporated age structure and observed per-capita birth and death rates, in order to estimate colonization levels following the introduction of the vaccine.

Two-serotype SIS model with Vaccination and Age Structure

Temime, Guillemot, and Boëlle [44] explored the effect of childhood vaccination on *S. pneumoniae* colonization and antimicrobial resistance through an age-structured two-serotype SIS model. Their main goals were to incorporate observed data, to investigate the epidemiological characteristics of *S. pneumoniae* in a vaccinated population, and to study the potential changes in the distribution of resistance levels among colonized subjects after the introduction of the vaccine. All *S. pneumoniae* serotypes were categorized into one of two

groups: vaccine-type (VT; serotypes 4, 6B, 9V, 14, 18C, 19F, and 23F) or non-vaccine-type (NVT; all other serotypes).

Model Structure

Temime et al. utilized a very complex model, which incorporated not only colonization status, age group, and vaccination status, but also included antibiotic use and ten different levels of antimicrobial resistance. A simplified version of their model, which does not include antibiotic use or antimicrobial resistance, is presented below to allow a more straightforward comparison with the other models described. This simplification resulted in a two-serotype SIS model with three age groups. Subjects were divided into young children (less than 2 years old), older children (2 - 15 years old), and adults (16 years or older), as indicated by a third subscript.

Transmission parameters β_{ij} were taken to be equal for both serotype groups. Therefore, the basic reproductive numbers for each serotype group, R_{0i}, calculated as a function of the infectious contact rates, average duration of colonization, and mortality, were taken to be equal for both serotype groups. This model did not allow dual colonization.

Young children were vaccinated at rate v, and moved from the unvaccinated compartments, N_{SS1}, N_{SI1}, or N_{IS1}, to the vaccinated compartments, N_{VS} and N_{VI}. These two vaccinated compartments did not incorporate age structure. Vaccinated children moved into the adult compartments (N_{SS3}, N_{SI3}, and N_{IS3}) at rate ω, after remaining in the vaccinated compartments for an average of 15 years. The vaccine was taken to be 100% effective against colonization with VT serotypes, and was assumed to have no effect on NVT serotypes.

In the simplified Temime et al. model, as described in {2.5}, age group was denoted by subscript $i = 1, 2, 3$. Parameters g_0 and g_3 were included for brevity of presentation, and were set equal to zero. $\delta_{ij} = 1$ when $i = j$, and 0 when $i \neq j$.

$$dN_{SSi}/dt = \delta_{i1}\mu_N N_3 + g_{i-1}N_{SS(i-1)} - (\lambda_{1i} + \lambda_{2i})N_{SSi} + \gamma N_{SIi} + \gamma N_{ISi} + \delta_{i3}\omega N_{VS} - (\delta_{i1}\nu + g_i + \mu_i)N_{SSi}$$
$$dN_{SIi}/dt = g_{i-1}N_{SI(i-1)} + \lambda_{2i}N_{SSi} - \gamma N_{SIi} + \delta_{i3}\omega N_{VI} - (\delta_{i1}\nu + g_i + \mu_i)N_{SIi}$$
$$dN_{ISi}/dt = g_{i-1}N_{IS(i-1)} + \lambda_{1i}N_{SSi} - \gamma N_{ISi} - (\nu + g_i + \mu_i)N_{ISi}$$

{2.5}

$$dN_{VS}/dt = \nu(N_{SS1} + N_{IS1}) + \gamma N_{VI} - \lambda_{2i}N_{VS} - (\omega + \mu_2)N_{VS}$$
$$dN_{VI}/dt = \nu N_{SI1} + \lambda_{2i}N_{VSi} - \gamma N_{VI} - (\omega + \mu_2)N_{VI}$$

where $\lambda_{1i} = \sum_{j=1}^{3}\left(\frac{\beta_{ij}N_{ISj}}{N_j}\right)$ and $\lambda_{2i} = \left(\frac{\beta_{i2}N_{VI}}{N_2}\right) + \sum_{j=1}^{3}\left(\frac{\beta_{ij}N_{SIj}}{N_j}\right)$

and $N_1 = N_{SS1} + N_{SI1} + N_{IS1}$
$N_2 = N_{SS2} + N_{SI2} + N_{IS2} + N_{VS} + N_{VI}$
$N_3 = N_{SS3} + N_{SI3} + N_{IS3}$

Equilibria

This model could be viewed as a more complex version of the Lipsitch model, with the same four theoretically possible outcomes (E_0, E_1, E_2, and E_{12}). Since dual colonization was not present in this model (effectively $c_i = 0$), coexistence at equilibrium (E_{12}) was only possible prior to vaccination if $R_{01} = R_{02}$, as shown in Figure 2.2. Multiple coexistence equilibria were present if $R_{0i} > 1$ and vaccination was not present; at those equilibria, $N_{SI\cdot} + N_{IS\cdot} = 1 - 1/R_0$. If $R_{01} \neq R_{02}$, the serotype with the larger basic reproductive number was present at equilibrium (i.e., E_1 if $R_{01} > R_{02}$, or E_2 if $R_{01} < R_{02}$).

If both serotypes were present at the pre-vaccination equilibrium, no dual-serotype equilibrium was stable after vaccination was introduced. In this situation, multiple equilibria were present, where $N_{SS\cdot} = 1/R_0$ and $N_{IS\cdot} + N_{SI\cdot} = 1 - 1/R_0$. Therefore, the ratio of the VT serotypes to the NVT serotypes at equilibrium depended on the initial conditions specified, given that there were multiple equilibria.

Outcomes

Model-based estimates of colonization prevalence before and after vaccination were not specifically provided in the Temime et al. paper, so we estimated those values from the graphs provided. Prior to vaccination, approximately 17% of subjects were colonized with VT serotypes, and approximately 7% of subjects colonized with NVT serotypes, resulting in 24% total colonization among the general population. The scenario incorporating 80% vaccination of young children estimated that VT colonization approached zero approximately ten years after the introduction of vaccination, while NVT colonization increased to approximately 22% during that time, and remained at approximately the same level for the next ten years.

This model was the first *S. pneumoniae* dynamic transmission model to incorporate observed data, in an effort to predict outcomes following the introduction of a pneumococcal conjugate vaccine. It incorporated differing mortality rates by age group, and also allowed infectious contact rates to vary between different pairs of age groups.

The focus of this model was to incorporate antibiotic use and antimicrobial resistance, and by simplifying the model, we adapted it to a use not originally intended by the authors. Certain aspects of the model, such as the assumption that the vaccine was 100% effective against colonization, were disproven in more recent research [31-39]. Transmission parameter values used in this model were selected to fit colonization patterns observed in France, and may not be applicable to other populations. Furthermore, most *S. pneumoniae* incidence and cost data available focus on IPD. Therefore, a model that incorporates both colonization and IPD is essential.

Two-serotype SIS model for colonization and IPD

Melegaro et al. [46] more recently developed an age-structured transmission model that incorporated both colonization and IPD. Colonization information was based on a longitudinal study conducted in the United Kingdom in 2001 and 2002. Unknown parameter

values were fitted using a second set of epidemiological data from England and Wales, and IPD surveillance data from the United States. The authors estimated IPD incidence in the United Kingdom for the twenty-year period following the introduction of the PCV7 vaccine.

Model Structure

This SIS colonization model used 100 different age cohorts, denoted by subscripts i = 0, 1, …, 99. Death only occurred in the oldest cohort of subjects, and total births and deaths were taken to be equal. Each cohort was the same size, with $1/100^{th}$ of the population in each cohort. Ageing was taken to occur at the end of each year, when subjects in the i^{th} age group moved to the $(i+1)^{th}$ age group.

Vaccination was similarly considered a discrete event, with all infants in a cohort assumed to be vaccinated at the same time, depending on the particular vaccination schedule used in the model. The effect of vaccination waned at rate ω, where $1/\omega$ was the average duration of protection. Vaccination provided partial protection e against colonization by VT serotypes, and no protection against colonization by NVT serotypes, effectively setting $k = 1$.

The 100 age cohorts were combined into six age groups (0-1, 2-4, 5-9, 10-19, 20-39, and 40-99 years, inclusive) before estimating the transmission parameters β_{ij}. As with the Temime et al. model, serotypes were grouped into VT and NVT serotypes, although the VT serotype group also included serotype 6A in this model. The transmission parameters were calculated assuming a combination of assortative (β^a) and proportionate (β^p) mixing between age groups, where $\beta_{ij} = \varepsilon \beta^a_{ij} + (1- \varepsilon) \beta^p_{ij}$ for each serotype group. This approach allowed the authors to estimate a reduced number of transmission parameters than is required by other methods. The authors used detailed information from the longitudinal colonization study, allowing them to calculate more transmission parameter values than is typically possible for other researchers. Average durations of colonization were taken to be the same for all serotypes, and varied by age group (0-1, 2-4, 5-17, and 18+ years).

The movement between compartments is described by the equations shown in {2.6}, where i = 0 to 99. Because birth, ageing and vaccination occur annually, terms corresponding to these processes do not appear in the differential equations.

Fractions of colonized subjects were then assumed to develop IPD. These fractions, or case:colonization ratios, were estimated separately for each serotype group for nine different age groups (<2 months, 2-11 months, 1-4, 5-14, 15-24, 25-44, 45-64, 65-74, and 75+ years). Vaccination was characterized as leading to a reduction in the chance of colonization by VT serotypes, but not having any further effect on the development of IPD.

$$
\begin{aligned}
dN_{SSi}/dt &= \gamma N_{SIi} + \gamma N_{ISi} - (\lambda_1 + \lambda_2) N_{SSi} - \nu_i N_{SSi} + \omega N_{VSSi} \\
dN_{SIi}/dt &= \lambda_2 N_{SSi} - c_1 \lambda_1 N_{SIi} - \gamma N_{SIi} + \gamma N_{IIi} - \nu_i N_{SIi} + \omega N_{VSIi} \\
dN_{ISi}/dt &= \lambda_1 N_{SSi} - c_2 \lambda_2 N_{ISi} - \gamma N_{ISi} + \gamma N_{IIi} - \nu_i N_{ISi} + \omega N_{VISi} \\
dN_{IIi}/dt &= c_1 \lambda_1 N_{SIi} + c_2 \lambda_2 N_{ISi} - 2\gamma N_{IIi} - \nu_i N_{IIi} + \omega N_{VIIi} \\
dN_{VSSi}/dt &= \gamma N_{VSIi} + \gamma N_{VISi} - (e\lambda_1 + \lambda_2) N_{VSSi} + \nu_i N_{VSSi} - \omega N_{VSSi} \\
dN_{VSIi}/dt &= \lambda_2 N_{VSSi} - ec_1 \lambda_1 N_{VSIi} - \gamma N_{SIi} + \gamma N_{VIIi} + \nu_i N_{VSIi} - \omega N_{VSIi} \\
dN_{VISi}/dt &= e\lambda_1 N_{VSSi} - c_2 \lambda_2 N_{VISi} - \gamma N_{ISi} + \gamma N_{VIIi} + \nu_i N_{VISi} - \omega N_{VISi} \\
dN_{VIIi1}/dt &= ec_1 \lambda_1 N_{VSIi} + c_2 \lambda_2 N_{VISi} - 2\gamma N_{VIIi} + \nu_i N_{VIIi} - \omega N_{VIIi}
\end{aligned}
\qquad \{2.6\}
$$

where

$$
\begin{aligned}
\lambda_{1i} &= \sum_j \beta_{1ij} (N_{ISj} + N_{VISj} + N_{IIj} + N_{VIIj}) \\
\lambda_{2i} &= \sum_j \beta_{2ij} (N_{SIj} + N_{VSIj} + N_{IIj} + N_{VIIj})
\end{aligned}
$$

Equilibria

This model could also be viewed as a more complex version of the Lipsitch model, where k = 1. We could not explicitly calculate the values of R_{0i} used in this model, due to the lack of information provided on transmission parameters and demographic information. Based on the graphs provided in their paper, we concluded that both serotype groups were present prior to vaccination (E_{12}), and that the VT serotype group eventually died out following vaccination (E_2).

Outcomes

Estimates of unknown parameters (c_2, ε, ω, e) were obtained by fitting the model to IPD incidence data from the United States. Competition parameters were estimated to be different between serotype groups, with a greater reduction in co-colonization caused by VT serotypes (c_1= 0.5) than by NVT serotypes (c_2 = 0.85). Vaccination was estimated to provide a 75.6% reduction in the chance of colonization by VT serotypes among vaccinated subjects (i.e., the direct effect of vaccination), with an average duration of protection of 8.3 years. These parameter values were estimated by fitting model outputs to observed data. This one-stage characterization of vaccination effects as causing a large reduction in VT colonization, and no additional reduction in IPD, differs from the two-stage vaccination effect incorporated into the Snedecor et al. model, and our SIRS model, as discussed below.

Transmission parameters were estimated as 87% assortative and 13% proportionate, meaning that 13% of interactions were distributed across the five age groups according to proportionate mixing, and the other 87% were interactions within a subject's own age group. Average durations of colonization were 72 days (0-1 year olds), 28 days (2-4 year olds), 18 days (5-17 year olds), or 17 days (18+ year olds). Case:colonization ratio estimates for the development of IPD ranged from 0.00001 to 0.00064, depending on the serotype and age groups. Case:colonization ratios were equal between serotypes for 1-14 year olds, but ratios for NVT serotypes were approximately 2-3 times as high as ratios for VT serotypes in the other age groups.

At equilibrium, this model estimated that the total IPD cases would decrease by 63% in children less than 5 years old, and by 35% in the rest of the population by the end of the first ten years. Based on the figure presented, this model estimated that NVT IPD incidence would increase by approximately 20% during the first ten years after the vaccine was introduced in the United Kingdom, and that the total number of NVT IPD cases each year would remain approximately constant after that time. Similarly, the model also estimated

that VT IPD incidence would approach zero approximately ten years after the introduction of the vaccine in the United Kingdom. Since IPD incidence was calculated as a fraction of colonization incidence, this implied that NVT colonization incidence would increase, while VT serotypes would virtually disappear in that time period. Predicted outcomes from this model were very sensitive to changes in the protective effect of vaccination and the competition between serotypes.

These outcomes differed somewhat from the disease patterns observed in the United States in the ten years following the introduction of the pneumococcal conjugate vaccine. Based on CDC-reported IPD [17], annual IPD incidence decreased by 76% in children under 5 years old, (from 93.5 to 22.2 cases per 100,000 children) from 1999 to 2009. It was difficult to compare specific outcomes, due to the limited amount of publicly available information regarding IPD incidence by serotype group and the minimal amount of numeric outcomes provided by Melegaro et al.

Certain parameter values used in this model differed from the values indicated in more recent literature. Other research indicated that the vaccine was associated with a 29%-69% reduction in colonization [31; 32; 35-39], and an approximate 95% total decline in VT IPD incidence [68; 69], in comparison with this model's characterization of vaccine effects as leading to a 75.6% reduction in VT colonization with no additional decrease in the chance of developing VT IPD. Other authors have suggested that competition between serotypes results in a 20% decrease in co-colonization among already colonized individuals (c_i = 0.80) [70]. A more recent publication by Melegaro et al. [71] stated that the United Kingdom was experiencing higher serotype replacement than has been observed in the United States, suggesting that competition parameters vary between the two countries. Furthermore, case:colonization ratio estimates in this model were calculated for England and Wales; these may vary slightly from case:colonization ratios for the United States. In particular, these case:colonization ratios calculated by Melegaro et al. were two to three times higher for NVT

serotypes than VT serotypes among most age groups. Therefore, this model may not provide the best fit for estimating colonization and IPD in the United States.

Single-serotype SIS model with colonization and IPD

Snedecor et al. [47] recently developed a single-serotype, age-structured dynamic transmission model for *S. pneumoniae* which included both colonization and IPD. The purpose of this model was to quantify the direct and indirect benefits of infant vaccination and simulate the acquisition of colonization and IPD in the overall population during the ten years following the introduction of the vaccine in the United States. Unlike previous models, this model did not differentiate between VT and NVT serotypes. However, it was the first model to incorporate a catch-up vaccination period, to allow slightly older children to be vaccinated following the introduction of the vaccine, as occurred in the United States.

Model Structure

In this model, a fraction of subjects who became colonized immediately developed IPD, and a fraction of those developing IPD immediately died. IPD was divided into three diseases: meningitis, bacteremia, and invasive pneumonia. Transmission parameters β_i were estimated from disease incidence [72], colonization incidence [5], and average duration of colonization [9]. Pre-vaccination estimated IPD incidence was presented for eight age groups, average colonization levels were presented for four age groups, and average durations of colonization were presented for six age groups. However, the specific methods used for parameter estimation were not discussed. Subjects with IPD were not considered infectious subjects when determining the transmission of infection, although this is expected to have a very small effect due to the small fractions of subjects developing IPD.

Subjects were born into the youngest age group (less than 2 years) at a constant birth rate, and matured into the older age groups (2-4, 5-17, 18-49, 50-64, 65+ years) at constant rates g_i (i = 1-5). Subjects in each of the six age groups died from non-IPD related death at rates μ_i (i = 1-6). Additional deaths were caused by IPD at age-group- and disease-specific rates μ_{Xi},

where $X = M, B, P$ indicates meningitis, bacteremia, or pneumonia respectively. Total birth and death rates were taken to be equal, to ensure a constant population size. The authors did not specify how disease-related deaths were reconciled with the assumption of constant population size; we assume that they did this by setting birth rates equal to the sum of IPD-related and unrelated death rates, and allowing birth rates to vary over time.

The fractions of subjects in the two youngest age groups who were vaccinated during the first five years of vaccination varied, to approximate patterns in the United States following the introduction of the vaccine. In the fifth and subsequent years, 95% of children under 2 years old and 0% of children 2-4 years old were vaccinated. Vaccine protection was considered to remain constant for the ten-year duration of the model. Snedecor et al. incorrectly labeled fractions of subjects vaccinated each year as the "rate of vaccine administration at year *t*." Although it was unclear whether the fractions of subjects vaccinated were incorrectly substituted for per-capita vaccination rates, labeled $v_1(t)$ and $v_2(t)$ in the equations shown in {2.7}, we assume that those authors incorporated vaccination rates necessary to achieve the observed vaccination levels among children in those age groups.

Since VT and NVT serotypes were not differentiated in this model, the vaccine was characterized as providing 40% protection against all *S. pneumoniae* colonization ($e = 0.60$), and an additional 90% protection against IPD ($r = 0.10$). The characterization of vaccine effect against colonization was based the reduction in VT colonization among vaccinated children in Dagan et al. [32]. This two-stage characterization of vaccine effects effectively characterized the vaccine as having 94% efficacy against all IPD, including that caused by serotypes not included in the vaccine.

This model characterized birth, death, and growth between age classes as occurring only among susceptible and colonized subjects. The average duration of colonization varied by age, with average durations of 2 months for children less than 2 years old, 1.5 months for children 2-4 years, and 0.75 months for all other subjects, based on Ekdahl et al.'s study of

the duration of pneumococcal carriage. Average durations of IPD were assumed to be 1 month for all disease types and age groups, due to lack of available information.

The model equations shown in {2.7} characterize the movement between compartments in this model, where N_{Si} and N_{VSi} include susceptible subjects, N_{Ii} and N_{VIi} contain colonized subjects, and N_{Mi}, N_{Bi}, N_{Pi}, N_{VMi}, N_{VBi}, and N_{VPi} include subjects with IPD. Colonized subjects developed IPD at age-group- and disease-specific rates ζ_{Xi}, and recover from IPD at rates η_{Xi}. A fraction d_{Xi} of subjects who developed IPD immediately became case fatalities. Therefore, a fraction of the subjects in the N_{SIi} compartment became colonized at rate λ_i, (1-(ζ_{Mi}+ ζ_{Bi}+ ζ_{Pi})) of those subjects remained in the colonized compartment for an average duration of $1/\gamma$ time units, ($d_{Mi}\zeta_{Mi}$+d_{Bi} ζ_{Bi}+ $d_{Pi}\zeta_{Pi}$) died immediately from fatal IPD, and ((1-d_{Mi})ζ_{Mi}+(1-d_{Bi})ζ_{Bi}+ (1-d_{Pi})ζ_{Pi}) remained in an IPD compartment for an average duration of $1/\eta_{Xi}$ time units before becoming susceptible again.

In previously described models, the forces of infection, λ_i, were calculated as the products of the transmission parameters (β_i) and the numbers of colonized subjects as shown in equations {2.3}, {2.4}, {2.5}, and {2.6}. However, the structural forms of the forces of infection used in this model varied by age group, as shown in {2.7}. The forces of infection for the youngest three age groups (children less than 18 years of age) were consistent with earlier models. The forces of infection for 18 – 64 year olds were only dependent on the colonized subjects in the particular age group (18-49 or 50-64 years) and the number of colonized subjects less than 2 years old. The force of infection for subjects at least 65 years of age, λ_6, was dependent on the number of colonized subjects in that age group, and the square of the number of colonized subjects less than 2 years old. The authors did not fully explain this apparently ad hoc characterization of forces of infection, but stated that it was due to influences between different age groups when approximating observed data.

$$
\begin{aligned}
dN_{Si}/dt &= \delta_{i1}\mu_B + g_{i-1}N_{S(i-1)} + \gamma N_{Ii} - \lambda_i N_{Si} + \eta_{Mi}N_{Mi} + \eta_{Bi}N_{Bi} + \eta_{Pi}N_{Pi} - (\delta_{i1}v_1(t) + \delta_{i2}v_2(t))N_{Si} - (g_i + \mu_i)N_{Si} \\
dN_{Ii}/dt &= g_i N_{I(i-1)} - \gamma N_{Ii} + \lambda_i N_{Si} - (\zeta_{Mi} + \zeta_{Bi} + \zeta_{Pi})\lambda_i N_{Si} - (\delta_{i1}v_1(t) + \delta_{i2}v_2(t))N_{Ii} - (g_i + \mu_i)N_{Ii} \\
dN_{Mi}/dt &= (1 - d_{Mi})\zeta_{Mi}\lambda_i N_{Si} - \eta_{Mi}N_{Mi} \\
dN_{Bi}/dt &= (1 - d_{Bi})\zeta_{Bi}\lambda_i N_{Si} - \eta_{Bi}N_{Bi} \\
dN_{Pi}/dt &= (1 - d_{Pi})\zeta_{Pi}\lambda_i N_{Si} - \eta_{Pi}N_{Pi} \\
dN_{VSi}/dt &= (\delta_{i1}v_1(t) + \delta_{i2}v_2(t))N_{Si} + g_{i-1}N_{VS(i-1)} + \gamma N_{VIi} - e\lambda_i N_{VSi} + \eta_{Mi}N_{VMi} + \eta_{Bi}N_{VBi} + \eta_{Pi}N_{VPi} - (g_i + \mu_i)N_{VSi} \\
dN_{VIi}/dt &= (\delta_{i1}v_1(t) + \delta_{i2}v_2(t))N_{Ii} + g_{i-1}N_{VI(i-1)} - \gamma N_{VIi} + e\lambda_i N_{VSi} - er(\zeta_{Mi} + \zeta_{Bi} + \zeta_{Pi})\lambda_i - (g_i + \mu_i)N_{VIi} \\
dN_{VMi}/dt &= er(1 - d_{Mi})\zeta_{Mi}\lambda_i N_{Si} - \eta_{Mi}N_{Mi} \\
dN_{VBi}/dt &= er(1 - d_{Bi})\zeta_{Bi}\lambda_i N_{Si} - \eta_{Bi}N_{Bi} \\
dN_{VPi}/dt &= er(1 - d_{Pi})\zeta_{Pi}\lambda_i N_{Si} - \eta_{Pi}N_{Pi}
\end{aligned}
\qquad \{2.7\}
$$

$$
\lambda_i = \beta_i \left[\varepsilon_i (N_{Ii} + N_{VIi}) + (1 - \varepsilon_i) \sum_{j=1}^{6} (N_{Ij} + N_{VIj}) \right] \qquad \text{for } i = 1, 2, 3
$$

where

$$
\lambda_i = \beta_i \left[\varepsilon_i (N_{Ii} + N_{VIi}) + (1 - \varepsilon_i)(N_{I1} + N_{VI1}) \right] \qquad \text{for } i = 4, 5
$$

$$
\lambda_6 = \beta_6 \left[\varepsilon_6 (N_{I6} + N_{VI6}) + (1 - \varepsilon_6)(N_{I1} + N_{VI1})^2 \right]
$$

Equilibria

The colonization portion of this model was an SIS model. In order to explore the stability of the model, we simplified it to a single age group and a single IPD group, and assumed $\lambda = \beta(N_I + N_{VI})$. In the absence of IPD ($\zeta_X = 0$), this model was stable at the disease-free equilibrium, where no infection was present, or at the endemic equilibrium, where $1/R_0$ subjects were susceptible and $1\text{-}1/R_0$ subjects were colonized. In the presence of IPD, $1/(R_0(1\text{-}\zeta_X))$ were susceptible and $(\zeta_X(\mu+\gamma)(1\text{-}d)/(\eta(1\text{-}\zeta_X)))$ N_I had IPD, and the remainder of the population was colonized.

We then expanded these results to the complete model. If the average number of secondary colonizations caused by a colonized individual was less than one, the system would converge to the disease-free equilibrium, with fractions of subjects in different age groups dependent

primarily on g_i and μ_i. The age-specific IPD rates ζ_{Xi} and associated IPD-related death rates d_{Xi} also affected the fractions of subjects in different age groups. After the introduction of the vaccine, the ratio of subjects in N_{Si} and N_{VSi} depended on $v_1(t)$ and $v_2(t)$. The fractions of colonized and IPD subjects within each age group then depended on β_i, ε_i, ζ_{Xi}, d_{Xi}, and η_{Xi}.

Outcomes

No colonization outcomes were presented by the authors. This model estimated that IPD decreased significantly within ten years after the introduction of the vaccine. Children 0-4 years of age were estimated to have an 85% reduction in IPD, with smaller reductions in older age groups (5-17 years, 66%; 18-49 years, 46%; 50-64 years, 31%; 65+ years, 49%). These estimates of reduction in IPD were slightly higher than were actually observed (<2 years, 81%; 2-4 years, 59%; 5-17 years, 20%; 18-49 years, 40%; 50-64 years, 12%; 65+ years, 37%) in the United States [17].

This model exhibited a better fit to observed IPD levels in the United States following the introduction of the vaccine than was shown by the Melegaro et al. model. It incorporated vaccination patterns that approximate those observed in the United States immediately following the introduction of the vaccine, including catch-up vaccination of older children. It also incorporated a two-stage vaccination effect, with vaccination leading to decreased colonization and additional decreases in IPD.

However, there were certain parameter estimates and aspects of model structure that did not reflect the most current findings. This model's characterization of the vaccine as having 94% efficacy against all IPD was approximately similar to the decrease in total IPD found during a study conducted in 2002-2003, which found an 84-94% decrease, depending on age group, when compared to a four year period prior to the introduction of the vaccine [21]. This study found a 99-100% decrease in VT IPD and a 41% decrease to 21% increase in NVT IPD, depending on age group, during that period. More recent research, comparing 2007 IPD incidence to the incidence in 1998-1999, found a 45% decrease in IPD in the entire

population, with a 94% decrease in VT IPD [73]. This implied that IPD due to NVT serotypes (including all serotypes not in the vaccine) increased by 40% overall, and by 38% among children less than five years old, during that time. Therefore, the assumption that vaccination provided protection against all IPD for the ten-year period used in the model was not supported by observed data.

This model characterized vaccination as affecting all serotypes, not just the seven serotypes included in the vaccine. Forces of infection (λ_i) did not include subjects with IPD, and were calculated differently for different age groups. The intent of this model was to predict changes in IPD during the first ten years following the introduction of the vaccine. Therefore, a different model is needed to estimate changes in IPD after the first ten years.

Conclusion

These models provided the framework necessary to explore the interactions of two serotypes, or two groups of serotypes, of *S. pneumoniae* bacteria. The Lipsitch and Zhang et al. models introduced the initial theoretical framework. The Temime et al. model expanded that framework to a real-life setting, and incorporated observed data to provide estimates of colonization incidence following the introduction of the vaccine. The Melegaro et al. model incorporated both colonization and IPD, but its estimates did not reflect the pattern of IPD observed in the United States following the introduction of the vaccine. The Snedecor et al. model provided a better fit to observed IPD data, but characterized vaccination as providing permanent protection against all serotypes, making it a poor candidate for long-term model predictions.

However, the current models did not exhibit the best possible fit to the observed data. The Temime et al. model characterized vaccination as providing complete protection against colonization by VT serotypes, and the Melegaro et al. model used a one-stage vaccine effect, which over-estimated the observed reduction in VT colonization among vaccinated subjects [31-39] and underestimated the reduction in VT IPD among vaccinated subjects [68; 69].

The Snedecor et al. model characterized vaccination as causing reductions in colonization and IPD due to all serotypes, with no waning of vaccine protection, due to the model's intended duration of ten years. While each of these models was a valuable contribution, none of these models both provided a good fit to observed data in the United States, and was suitable for estimating colonization and disease patterns more than ten years after the introduction of the vaccine.

In the following chapter, we propose an age-structured dynamic transmission model that incorporated recently published research to provide estimates of both colonization and IPD following the introduction of the vaccine in the United States.

3. AGE-STRUCTURED SIRS DYNAMIC TRANSMISSION MODEL FOR *S. PNEUMONIAE*

Introduction

Streptococcus pneumoniae is a leading cause of pneumonia, bacteremia, and meningitis, and is a considerable public health concern worldwide. This bacterium has been a focus of vaccine development over the last two decades. When the bacteria adhere to the epithelial lining of the upper respiratory tract, it is considered *colonization* [2]. If the bacteria penetrate the mucosal barrier and enter the tissues, it is considered pneumococcal *infection* [3]. Pneumococcal disease occurs when the body's immune system reacts to the presence of the bacteria, such as the development of fluid in the lungs (pneumonia). Invasive pneumococcal disease (IPD) occurs when bacteria enter a normally sterile site, such as the blood or cerebrospinal fluid.

Asymptomatically colonized subjects serve as the primary source of transmission to other hosts [3]. *S. pneumoniae* is considered one of many bacterial flora normally found in the upper respiratory tract of children, and is typically spread through exposure to nasal aspirate from a colonized individual, such as through sneezing or coughing. Exposure does not always lead to colonization [4].

The first pneumococcal conjugate vaccine (PCV7) was licensed for use in children in the United States in February 2000 [21]. It targets seven of the more than 90 serotypes identified to date. Serotypes were categorized as vaccine-type (VT) or non-vaccine-type (NVT) serotypes, based on the PCV7 vaccine. This vaccine has been shown to decrease colonization [31-39] and IPD [22; 24-30] caused by VT serotypes. The majority of studies have observed both a decrease in VT serotypes and a slight increase in NVT serotypes (due to *serotype replacement)*, in colonization [31; 33-36; 38-40] and IPD [22; 24; 26-28].

Earlier *S. pneumoniae* transmission models [42-44] have suggested scenarios in which colonization levels following vaccination may be similar to, or even exceed, pre-vaccination

levels, due to serotype replacement. More recent models [46-48], which used colonization levels to predict IPD levels, were either limited to young children [48] or did not include a period of temporary immunity following colonization [46; 47] as has been suggested by other research [74-76].

We developed a deterministic age-structured transmission model, based on colonization and IPD levels reported in published research and US surveillance data. This model was used to estimate unknown parameter values, such as transmission probabilities, as well as to estimate colonization and IPD incidence, by age group and serotype group, following the introduction of the vaccine. This model was also used to predict the long-term effect of vaccination on colonization and IPD levels, including serotype replacement, and to compare these predictions to those based on recent models. Furthermore, a sensitivity analysis was used to examine the range of possible outcomes and to determine which parameter values have the greatest impact on these predictions, to allow further research to focus on providing improved estimates of those parameters.

This paper is organized as follows: first, a summary of the structure of past and current models is presented in the model development section. The notation, movement between compartments, and derivation of parameter values in the current model is discussed next. Estimated parameter values, model-based estimates of incidence, and findings from the sensitivity analysis are presented in the Results section. The importance of these results, and comparisons to results found in other research, are included in the Discussion section. Appendix A presents the differential equations used in the model.

Methods

Model Development

The spread of *S. pneumoniae* colonization in a population has often been modeled using compartmental epidemiological models, such as the SIS or SIRS models. In these models, subjects are either susceptible and able to be colonized (S), colonized with *S. pneumoniae* (I),

or recovered and temporarily immune (R). Several *S. pneumoniae* transmission models have been developed in recent decades. The first of the models [42] provided a structural framework for epidemiological models, utilizing a two-serotype SIS model with vaccination. Other researchers [43] expanded that framework into a two-serotype SIRS model with vaccination, and explored the effect of the average duration of temporary immunity on colonization levels. The effects of competition (i.e., the decreased chance of becoming colonized, if already colonized with a serotype from the other serotype group) were also explored. Age-structured models were developed to better incorporate differences between different age groups [7; 44], such as the inverse relationship between age and average duration of colonization [9].

Later models combined the earlier framework with vaccination, colonization, and IPD information collected since the introduction of the PCV7 vaccine in the United States. Some of these models focused on transmission within families [7; 10; 77] or were limited to transmission among young children [48], making it difficult to apply these results on a national level.

The most recent models [46; 47] utilized an age-structured SIS approach to estimate both colonization and IPD incidence since the introduction of the vaccine. Both of these models incorporated contact rates between different groups by using proportionate and assortative mixing between age groups. Neither model included a period of temporary immunity following the clearance of colonization.

We created an age-structured SIRS model, based on the colonization and IPD information available to date. This model includes a temporarily immune period following colonization as per recent research [74-76]. Transmission of the bacteria within the model was assumed to occur through clearly specified patterns of interaction and transmission between age groups. Recently published estimates of parameter values were also incorporated [31-39; 70], and unknown parameter values estimated by fitting the model to observed data. Model-

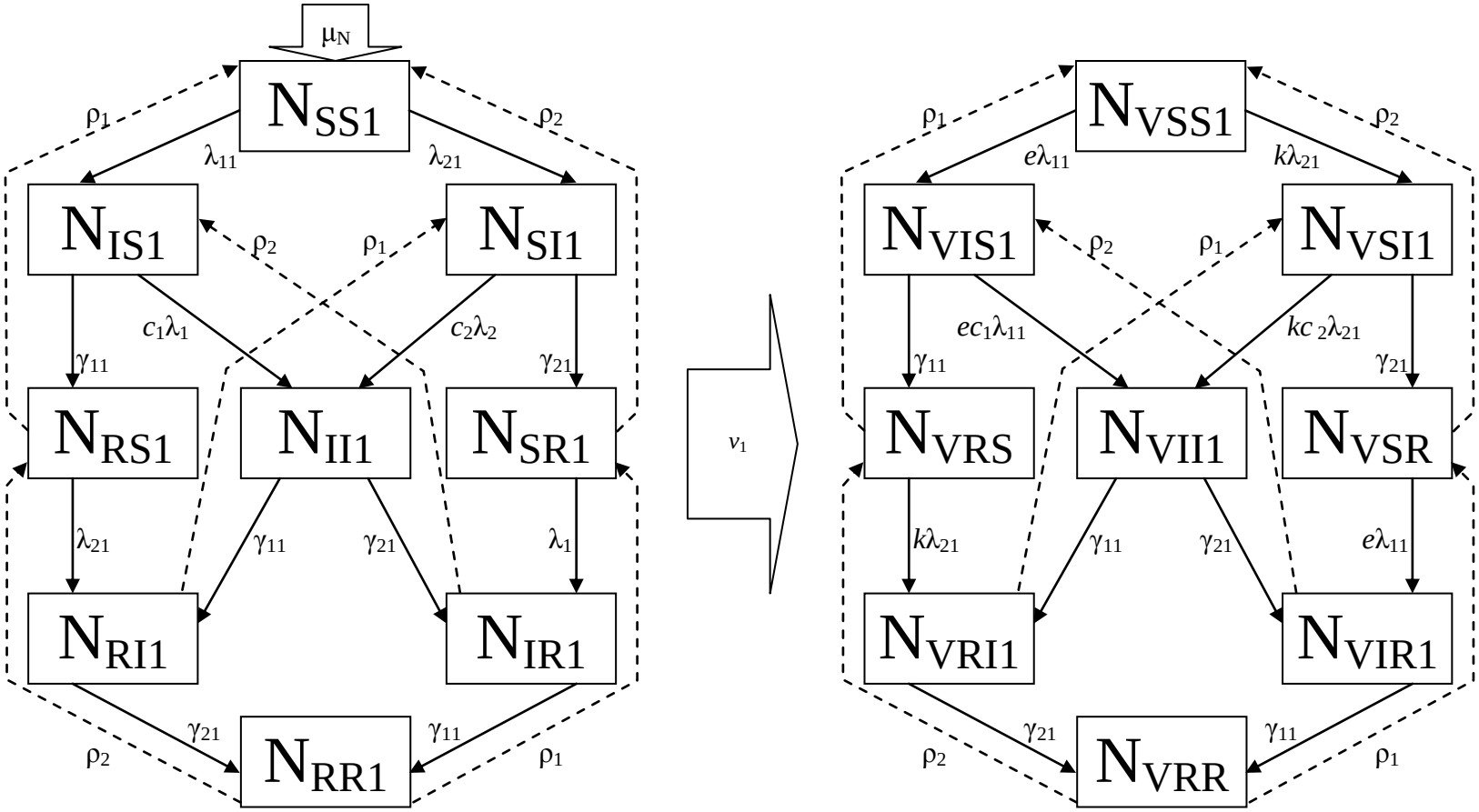
based estimates of both colonization and reported IPD incidence in the United States were then produced.

While the SIRS model can be used to model colonization in general, we were specifically interested in exploring the interactions of multiple serotypes of *S. pneumoniae*. This was particularly applicable since current vaccines are only effective against certain serotypes; for example, the PCV7 vaccine was targeted at seven serotypes: 4, 6B, 9V, 14, 18C, 19F, and 23F. Therefore, when one serotype, or group of serotypes, included in the vaccine was modeled concurrently with another serotype (or group of serotypes) not targeted by the vaccine, it is possible to explore the effects of the vaccine on the non-vaccine-type (NVT) serotype(s), as well as its effects on the vaccine-type (VT) serotype(s).

Model Structure

This model tracks the fractions of the population within each compartment. Movement between compartments was taken to be directly proportional to the number of subjects in the originating compartment, and is presented graphically in Figure 3.1 for the youngest age group. Model equations are displayed in Appendix A.

A two-serotype model was used to model this epidemiological system, in which the colonization statuses of subjects within a given population and age group were tracked with respect to two serotypes. Subjects were divided into three age groups: young children (less than 2 years), older children (2 - 15 years), and adults (16 years or older). These age groups were identical to those used in an earlier model [44], and allowed us to incorporate differing characteristics between age groups, such as transmission probabilities and average durations of colonization, while limiting the number of unknown parameter values being estimated. Each compartment was identified by subscripts indicating infection status, vaccination status, and age group. For example, a young child who was susceptible to both serotypes was included in N_{SS1}; an older child, colonized with the first serotype and recovered from the



Note: Subjects die at a rate of μ_1 and age in to the next age group at a rate of g_1, from every compartment.

Figure 3.1
Schematic of Colonization among Young Children Less Than 2 Years Old

second, was in N_{IR2}. Since vaccination was characterized as reducing, but not eliminating, colonization by VT serotypes, a set of compartments for vaccinated subjects was also included; similar subjects were included in N_{VSS1} and N_{VIR2}. The last subscript indicates the age group.

Invasive pneumococcal disease was modeled similarly, with subjects in compartments for VT and NVT IPD (N_{IPD1i} and N_{IPD2i}, respectively) for each age group. IPD cases were tracked separately from colonization cases within the model. Subjects developing IPD were counted concurrently in both the colonization and IPD compartments. The IPD compartments were therefore not counted when determining total population size, and did not feed back into the rest of the model for colonization or transmission. Therefore, IPD was characterized as not affecting transmission rates, or increasing death rates, among colonized subjects.

Movement Between Compartments

The flow of subjects within the youngest age group is shown in Figure 3.1. Movement within the other age groups was similar, with reduced vaccination among older children (v_2), and no vaccination among adults. Subjects were born into the youngest age group at rate μ_N. Within each compartment, these young children either died at rate μ_1, or aged into the next age group at rate g_1.

A fraction of subjects acquiring *S. pneumoniae* was considered to concurrently develop IPD, since most IPD occurs shortly after colonization [13]. The proportion of dually colonized individuals developing IPD was taken to be the sum of the proportions for each serotype.

Colonization Model and Parameter Estimation

Subjects were born susceptible, entering the N_{SS1} compartment. The population size was held constant, by taking the total birth rate to be equal to total death rate, whose value was based on the 2006 U.S. life expectancy of 77.7 years [78]. Subjects died at an age-group-specific rate (μ_i) within each compartment, based on age-specific 2006 U.S. death rates [78].

The average number of contacts with persons per day, by age group, was used to develop a matrix of contact rates. These contact rates were based on average contacts of subjects within fifteen different age groups in the United Kingdom [79]. Those fifteen age groups were collapsed into our three age groups using the following method. The total number of contacts made by a person was summed across appropriate age groups to calculate the total number of contacts who were in each of the three age groups. The weighted averages of the total number of contacts per day in each of the three age groups were then calculated. This resulted in the matrix of contact rates displayed in Table 3.1. These contact rates were included in the model as nine different parameters, m_{ij}, where i indicated the age group of the colonized subjects and j indicated the age group of the susceptible contacts.

Table 3.1
Average Number of Daily Contacts per Person, by Age Group.

Age of Colonized Subject (years)	Age of Contact (years)		
	<2	2 - 15	≥16
<2	0.768	2.648	5.464
2 - 15	0.323	6.834	6.346
≥16	0.149	1.568	8.879

A 3x3 matrix of transmission probabilities per contact between and within age groups was then created. Each transmission probability indicates the probability that a contact between a colonized subject and a susceptible subject results in the susceptible subject being colonized. In order to reduce the number of unique parameter values fitted, the probabilities were grouped into three parameters: transmission within young children (p_1); transmission between young children and older children, and within older children (p_2); and transmission between children and adults, and within adults (p_3). This type of approach is commonly implemented, to reduce the number of unknown parameters to a number that is able to be uniquely estimated from the data [80].

This results in a matrix of transmission probabilities of the form:

$$\begin{pmatrix} p_1 & p_2 & p_3 \\ p_2 & p_2 & p_3 \\ p_3 & p_3 & p_3 \end{pmatrix}$$

where columns and rows reflects age groups of the colonized subjects and contacts. VT and NVT transmission probabilities are likely to be different. Therefore, transmission probabilities for the NVT serotypes were taken to be a multiple d of the transmission probabilities for the VT serotypes. This maintained the same relationship between different age groups' transmission probabilities, within serotype group. Numbers of contacts between subjects in different age groups (i.e., the mixing matrix) were taken to be the same for both serotype groups. For each serotype, the matrix of transmission probabilities was combined with the average number of daily contacts to create a 3x3 matrix of transmission parameters (β_{ij}), commonly referred to as the 'who acquires infection from whom' (WAIFW) matrix [45].

Completely susceptible subjects became colonized at a rate λ_{ij}, called the force of infection, where i indicates the serotype (1=VT, 2=NVT) and j indicates the age group. This per-capita rate at which susceptible subjects were colonized, for a certain serotype and age group, was calculated as the sum over all three age groups of the products of the transmission parameters for colonization caused by each age group (e.g., β_{1j} for young children) and the fraction of subjects in the age group who were colonized by that serotype. These equations are shown in Appendix A.

Competition within each subject, i.e., the reduced chance that a subject already colonized with one serotype would become colonized with the other type, was accounted for by parameters c_1 and c_2. The force of infection λ_{ij} was multiplied by c_i, so that $c_i = 0.8$ indicated a 20% reduction in the rate of colonization among subjects who were already colonized [70].

In this model, subjects could not be dually colonized with serotypes from the same serotype (VT or NVT) group.

Average durations of colonization for young children (γ_{i1}^{-1}), where i indicates VT (i = 1) or NVT (i = 2) serotypes, were calculated using the weighted averages of durations of colonization for VT and NVT serotypes as recorded in 1996 - 2001 in Great Britain [81]. Since that research was limited to children under 2 years of age, the relationship between the average duration for VT colonization and NVT colonization in young children was used to calculate average durations for older subjects. Ratios of average durations of colonization for older children and adults, with respect to young children, were calculated as 70% and 46%, based on earlier research [9]. Information from these two sources was then used to calculate average durations of colonization for older age groups (γ_{i2}^{-1}, γ_{i3}^{-1}).

Fitting of Unknown Parameter Values (excluding Vaccination), Colonization Model

The transmission probabilities for VT serotypes (p_1, p_2, p_3), ratio of transmission probabilities between VT and NVT serotypes (d), and average durations of temporary immunity (ρ_1, ρ_2) were simultaneously fitted. These parameter values were fitted using observed data for five outcomes: pre-vaccination colonization levels by age group of 52.2%, 26.6%, and 8.5% from youngest to oldest age groups respectively [5], time to reinfection in young children of 14 weeks [82], and the fraction of colonization in young children caused by VT serotypes of 58% [81] at the pre-vaccination equilibrium. Average durations of temporary immunity, labeled ρ_1 for VT serotypes and ρ_2 for NVT serotypes, were set equal to each other.

The twenty-three parameter values based on earlier research were held constant, and values of the remaining five parameter values (p_1, p_2, p_3, ρ_i, d) which best fit the five colonization and time to reinfection characteristics recorded before the introduction of the vaccine were obtained using the Nelder-Mead simplex method [83], as implemented in the *fminsearch* function in Matlab v7.0. This method used unconstrained nonlinear optimization to find the

local minimum of the difference between recorded levels and model-based estimates of colonization levels by age group, time to reinfection in young children, and fraction of colonized young children having VT serotypes before vaccination. This was performed by minimizing

$$\sum_{i=1}^{5} \frac{(x_i - \hat{x}_i)^2}{x_i^2}$$

where x_i represents each of the five outcomes of interest (52%, 26.6%, and 8.5% colonization by age group; 14 week time to reinfection; 58% VT colonization) and $\hat{x}_i$ indicates the model-based estimate of that outcome. This optimization was performed using different sets of initial values to guard against the optimization routine finding a local minimum rather than the global minimum. The problems inherent in using five outcomes to fit five parameter values, such as the lack of standard errors associated with these estimates, are examined in the Discussion section of this chapter.

While the average durations of susceptibility varied by both age group and serotype, the average duration of temporary immunity (i.e., the time in the "R" compartment of the SIRS model) was taken to be constant for all subjects. This parameter value was determined as part of the optimization process, since the total time to reinfection is the sum of the average duration of temporary immunity and the average duration of susceptibility.

Vaccination and IPD: Model Structure and Fitting of Unknown Parameter Values

Vaccination was characterized as having a two-part effect on VT serotypes. First, it was taken to cause a reduction in the chance of colonization with VT serotypes, represented by the parameter *e*. Then, it was taken to cause an additional reduction in the chance of colonized subjects developing VT IPD, represented by the parameter *r*. The product of these two parameters represents the efficacy of the vaccine against VT IPD. This efficacy was taken to be 95% to reflect the 93.9% - 97.4% observed efficacy [68; 69].

Recent studies of the reduction in colonization of VT serotypes following vaccination have compared the prevalence of VT serotype colonization between vaccinated and unvaccinated subjects [31; 32; 35-39]. These studies found that vaccinated subjects had 29% - 69% lower colonization by VT serotypes than unvaccinated subjects in the same study. Due to the wide range of possible values, we chose to fit the parameter e to the available data, and subsequently calculated the value of r to ensure an overall 95% reduction in VT IPD among vaccinated subjects.

The parameter k accounted for the potential reduction in colonization by NVT serotypes among subjects who have been vaccinated. The value $k = 1.0$ was selected, representing no reduction in NVT colonization when vaccinated against VT serotypes, to see the most conservative (i.e., worst case) scenario. This value was chosen to allow the greatest possible increase in NVT colonization following the introduction of vaccination, since multiple studies [31; 33-36; 38-40] have observed such increases.

A second optimization was then performed to determine case:colonization ratios within each age group and serotype as well as the effect of the vaccine on VT colonization. A fraction f_{11} of unvaccinated young children who become colonized with VT serotypes was assumed to develop VT IPD. Among vaccinated young children, a smaller fraction ($r\ f_{11}$) who were colonized with VT serotypes developed VT IPD. A fraction f_{21} of young children who were colonized with NVT serotypes developed NVT IPD. Dually colonized individuals were allowed to develop both types of IPD.

The parameter values for six distinct case:colonization ratios, f_{ij}, and the reduction in the chance of VT colonization among vaccinated subjects were determined using the Nelder-Mead simplex method, as described previously. This optimization minimized the differences between reported annual IPD incidence in each year during 1999 – 2009 [17] and model-based estimates of IPD incidence for those eleven years, and concurrently minimized the differences between reported and predicted fractions of IPD within each age group which

were caused by VT serotypes in 1999 and 2007, based on serotype group and age group-specific IPD incidence published by Pilishvili et al. [73].

This optimization was performed by minimizing the objective function

$$\sum_{i=1}^{3} \sum_{j=1999}^{2009} \frac{(x_{ij} - \hat{x}_{ij})^2}{x_{ij}^2} + \sum_{i=1}^{3} \frac{(y_i - \hat{y}_i)^2}{y_i^2} + \sum_{i=1}^{3} \frac{(z_i - \hat{z}_i)^2}{z_i^2}$$

where i indicates age group, j indicates year, x_{ij} represents the reported IPD (VT and NVT combined) for that age group and year, and $\hat{x}_{ij}$ represents the model-based estimate of IPD for that age group and year. Similarly, y_i indicates the fraction of reported IPD due to VT serotypes in 1999, z_i indicates that fraction in 2007, and $\hat{y}_i$ and $\hat{z}_i$ represent the associated model-based estimates of those fractions. Model-based incidence was calculated by tracking the fractions of subjects within each age group who acquired IPD during a 52-week period, and calculating the number of cases per 100,000 subjects for that period. Since the incidence, not prevalence, of IPD was tracked in this model, and subjects with IPD were not taken to be more infectious than colonized subjects who had not developed IPD, the duration of IPD is irrelevant to this model.

While vaccination may reduce colonization by NVT serotypes (k), this model did not include an additional parameter to represent the potentially reduced rate of NVT IPD development among vaccinated subjects colonized with NVT serotypes. No information was available to suggest that such an effect is present.

Vaccination

Young children received vaccine at rate v_1, and moved from N_{xx1} to N_{Vxx1}. Annual vaccination levels in the youngest age group were based on the proportion of children less than twenty-four months of age who received one or more doses of either PCV7 or a pneumococcal vaccine of unknown type in a given birth cohort, calculated from the National Immunization Survey data for 2001 – 2009 [84].

Children who were 24 – 35 months of age received vaccination at rate v_2, and similarly moved from N_{xx2} to N_{Vxx2}. Annual vaccination levels in older children were similarly based on the calculated proportion of children who received their first dose of PCV7 or an unknown pneumococcal vaccine of unknown type at 24 - 35 months of age.

Where available, vaccination information based on actual birth cohort was used [85]. Vaccination by age group among other birth cohorts was calculated from weighted averages of the results for each survey year. The per-capita vaccination rate among young children was modeled as a time varying piecewise linear function for 2000-2001, and for each year thereafter, and taken to be constant from 2009 onward. In other words, each year (or pair of years, for 2000-2001), was fit separately, and the function constrained to be continuous. The per-capita vaccination rates among children 24 - 35 months of age (v_2) were similarly calculated through 2003, and taken to be constant after that time. The effect of vaccination was taken to wane when subjects move into the oldest age group. This was implemented by having both vaccinated and unvaccinated older children move into unvaccinated compartments as adults.

Results

Fitted Parameter Values

Five unknown parameter values were fitted by determining the values whose resulting model-based estimates best fit the total colonization levels by age group, the ratio of VT and NVT serotypes among colonized young children, and the mean time to recolonization in young children before the introduction of the vaccine that were reported in literature. Fitted values of transmission probabilities were $p_1 = 0.022$, $p_2 = 0.0013$, and $p_3 = 0.00031$ contact^{-1} for VT serotypes. Simultaneously, fitted values of NVT transmission probabilities were 160% of those for VT serotypes, resulting in equivalent transmission probabilities of 0.035, 0.0021, and 0.00050, respectively. The average duration of temporary immunity was

estimated to be 3.4 weeks ($\rho_1 = \rho_2 = 0.29$) for all subjects. The parameter values used to model colonization are shown in Table 3.2.

Table 3.2
Parameter Values Used to Model Colonization Prior to Vaccination

Parameter	Symbol	Value	Reference
Birth rate	μ_N	0.0002475 week $^{-1}$	[86]
Death rate, <2 yrs	μ_1	0.0006914 week $^{-1}$	[78]
Death rate, 2-15 yrs	μ_2	0.0000415 week $^{-1}$	[78]
Death rate, ≥16 yrs	μ_3	0.000245 week $^{-1}$	[78]
Average weekly contact rates			Based on
Among <2 yrs	m_{11}	5.38 week^{-1}	[79]
From <2 yrs to 2-15 yrs	m_{12}	18.54 week^{-1}	
From <2 yrs to ≥16 yrs	m_{13}	38.25 week^{-1}	
Among 2-15 yrs	m_{22}	47.84 week^{-1}	
From 2-15 yrs to <2 yrs	m_{21}	2.26 week^{-1}	
From 2-15 yrs to ≥16 yrs	m_{23}	44.42 week^{-1}	
Among ≥16 yrs	m_{33}	62.15 week^{-1}	
From ≥16 yrs to <2 yrs	m_{31}	1.04 week^{-1}	
From ≥16 yrs to 2-15 yrs	m_{32}	10.97 week^{-1}	
Average duration of colonization			Based on
Among <2 yrs, VT	γ_{11}^{-1}	16.0 weeks	[9; 81]
Among 2-15 yrs, VT	γ_{12}^{-1}	11.2 weeks	
Among ≥16 yrs, VT	γ_{13}^{-1}	7.3 weeks	
Among <2 yrs, NVT	γ_{21}^{-1}	10.3 weeks	
Among 2-15 yrs, NVT	γ_{22}^{-1}	7.2 weeks	
Among ≥16 yrs, NVT	γ_{23}^{-1}	4.7 weeks	
Average duration of temporary immunity, VT	ρ_1^{-1}	3.4 weeks	Fitted
Average duration of temporary immunity, NVT	ρ_2^{-1}	3.4 weeks	Fitted
Competition within subject, VT	c_1	0.8	[70]
Competition within subject, NVT	c_2	0.8	[70]
Transmission probability, VT			Fitted
Among <2 yrs	p_1	0.022 contact $^{-1}$	
Among 2-15 yrs and between <2 yrs and 2-15 yrs	p_2	0.0013 contact $^{-1}$	
Among ≥16 yrs and btwn. ≥16 yrs and other groups	p_3	0.00031 contact $^{-1}$	
Ratio of NVT:VT serotype transmission probability	d	1.6	Fitted

While this method of fitting unknown parameter values had the advantage of using observed data (instead of choosing parameter values that seem plausible), it is not a defensible statistical approach. This represented an overfitting of the data; for example, it is possible to exactly fit an n^{th} degree polynomial model to n+1 data points. However, the alternative was

to choose values for one or more of these parameters, or make additional assumptions, such as assuming $p_1 = p_2$.

Since we fit five parameter values using five quantities, we did not obtain an estimate of the standard error of each estimate. Therefore, we could not determine whether any of these parameter values were significantly different from zero (such as the transmission probabilities and average duration of temporary immunity) or significantly different from one (e.g., the ratio of transmission probabilities between serotype groups). Similarly, we could not determine if the pairwise differences between different transmission probabilities were significantly different from zero. The transmission probabilities did, however, exhibit the expected pattern, with highest transmission probability among interactions between children less than 2 years old, and the lowest transmission probability among interactions with adults. The ratio of transmission probabilities between serotype groups and average duration of temporary immunity both seemed to be reasonable, in that they were not extremely high or extremely low.

Seven additional parameter values were fitted by determining the values which, when combined with model-based estimates of colonization within each serotype and age group,

best fit the reported IPD incidence by age group and fractions of IPD due to VT serotypes. The fitted value for the effect of vaccination on VT colonization was 0.45, representing a 55% reduction in the chance of colonization. This was consistent with the 29% - 69% difference in colonization by VT serotypes between vaccinated and unvaccinated subjects observed in multiple studies [31; 32; 35-39]. The further effect of vaccination on the development of IPD, *r*, was then calculated to be 0.11, representing an 89% reduction in the chance of developing VT IPD among vaccinated subjects who were colonized with VT serotypes.

Fitted values for the case:colonization ratios (i.e., the proportion of colonized subjects who develop IPD) were 0.00017, 0.000088, and 0.00035, from youngest to oldest age group

respectively, for VT serotypes. Similarly, the fitted values for case:colonization ratios for subjects developing NVT IPD were 0.000022, 0.000049, and 0.00034 of subjects colonized with NVT serotypes. The parameter values used to model vaccination and IPD are shown in Table 3.3.

The next generation matrix [87] was calculated by multiplying the average duration of colonization and the transmission parameters for the appropriate age group and serotype. For each serotype group, the basic reproductive number, R_{0i}, was calculated as the maximum eigenvalue of each next generation matrix. The basic reproductive number for VT serotypes, R_{01}, was estimated by the model to be 1.90, and the basic reproductive number for NVT serotypes, R_{02}, was similarly estimated at 1.96.

Model Predictions

Model-based estimates of colonization prevalence and IPD incidence in 1999, 2010, and 2025 are presented in Table 3.4. Colonization prevalence is based on the middle of the year, and IPD incidence reflects the total number of new cases throughout the entire year. Estimated colonization prevalence is displayed by age group in Figure 3.2, and estimated IPD incidence is shown by age group in Figure 3.3. Model-based estimates of both colonization prevalence and IPD incidence, across all age groups, are shown in Figure 3.4.

Table 3.3
Additional Parameter Values Used to Model Vaccination and IPD

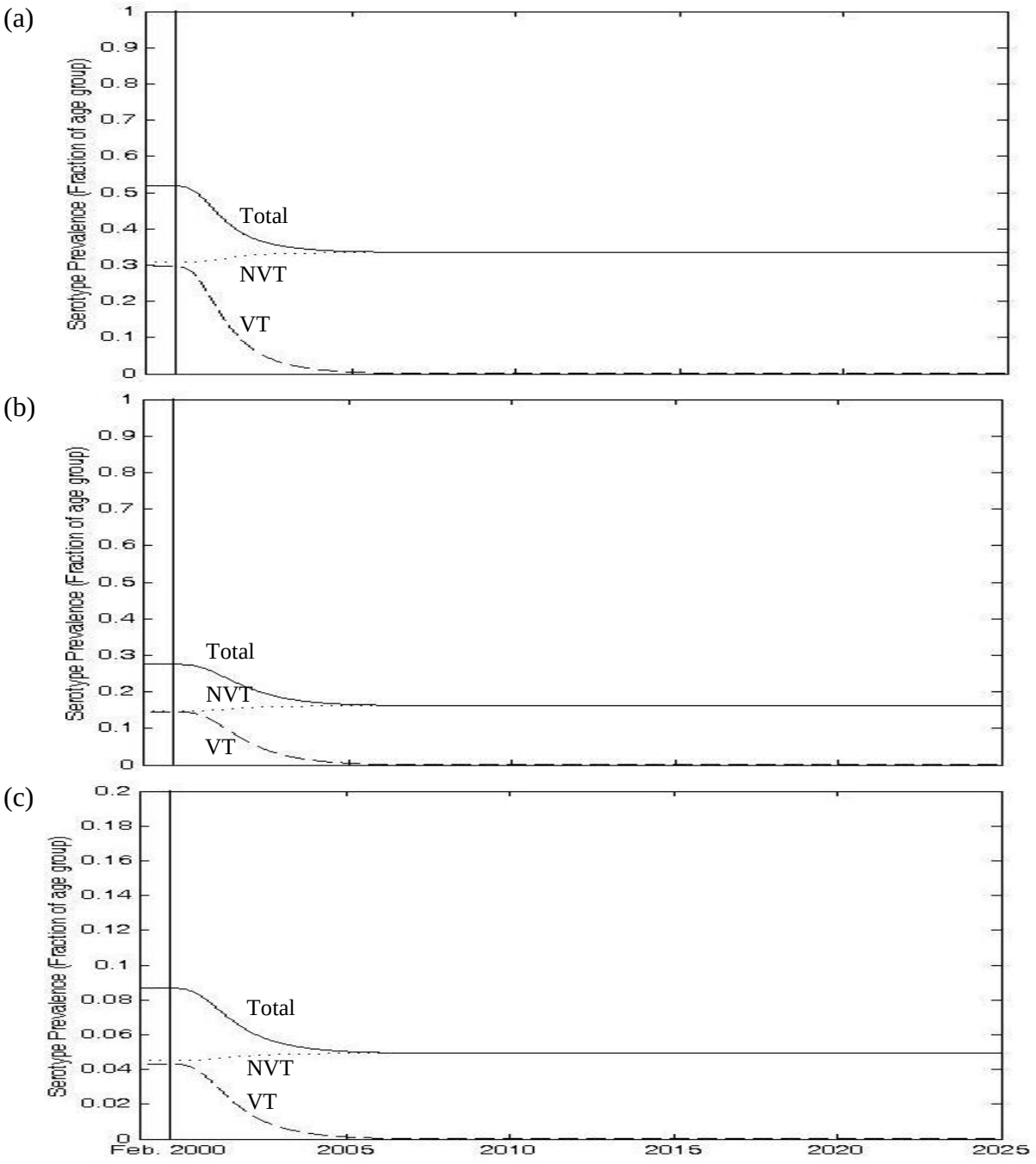
Parameter	Symbol	Value	Reference
Effect of vaccine on colonization by VT serotype	e	0.45	Fitted
Effect of vaccine on VT IPD	r	0.11	[68; 69]
Effect of vaccine on colonization by NVT serotype	k	1.0	--
Fraction of subjects being colonized who develop IPD, VT			Fitted
Among <2 yrs	f_{11}	0.0017	
Among 2-15 yrs	f_{12}	0.000088	
Among ≥16 yrs	f_{13}	0.00035	
Fraction of subjects being colonized who develop IPD, NVT			Fitted
Among <2 yrs	f_{21}	0.000022	
Among 2-15 yrs	f_{22}	0.000049	
Among ≥16 yrs	f_{23}	0.00034	

The model estimated that overall colonization levels in 2010 decreased from pre-vaccination levels in all age groups, to 33.5%, 16.2%, and 4.9% of subjects from youngest to oldest age group, respectively. This reflected the reduction in VT colonization due to 96% of young children and 47% of older children having received one or more doses of vaccine by this time. In addition to the direct reduction in colonization associated with vaccination, these levels reflected the effect of indirect reduction in colonization due to herd immunity, resulting in reductions in colonization levels among unvaccinated individuals, such as adults. Furthermore, these reductions in colonization reflect both decreases in VT colonization and slight increases in NVT colonization, where present. Young children showed the largest model-estimated changes, with VT colonization dropping from 29.5% to less than 0.1% of subjects, and NVT colonization increasing from 30.7% to 33.5% of subjects. The other age

Table 3.4
Model-Based Estimates of Colonization Prevalence and IPD Incidence

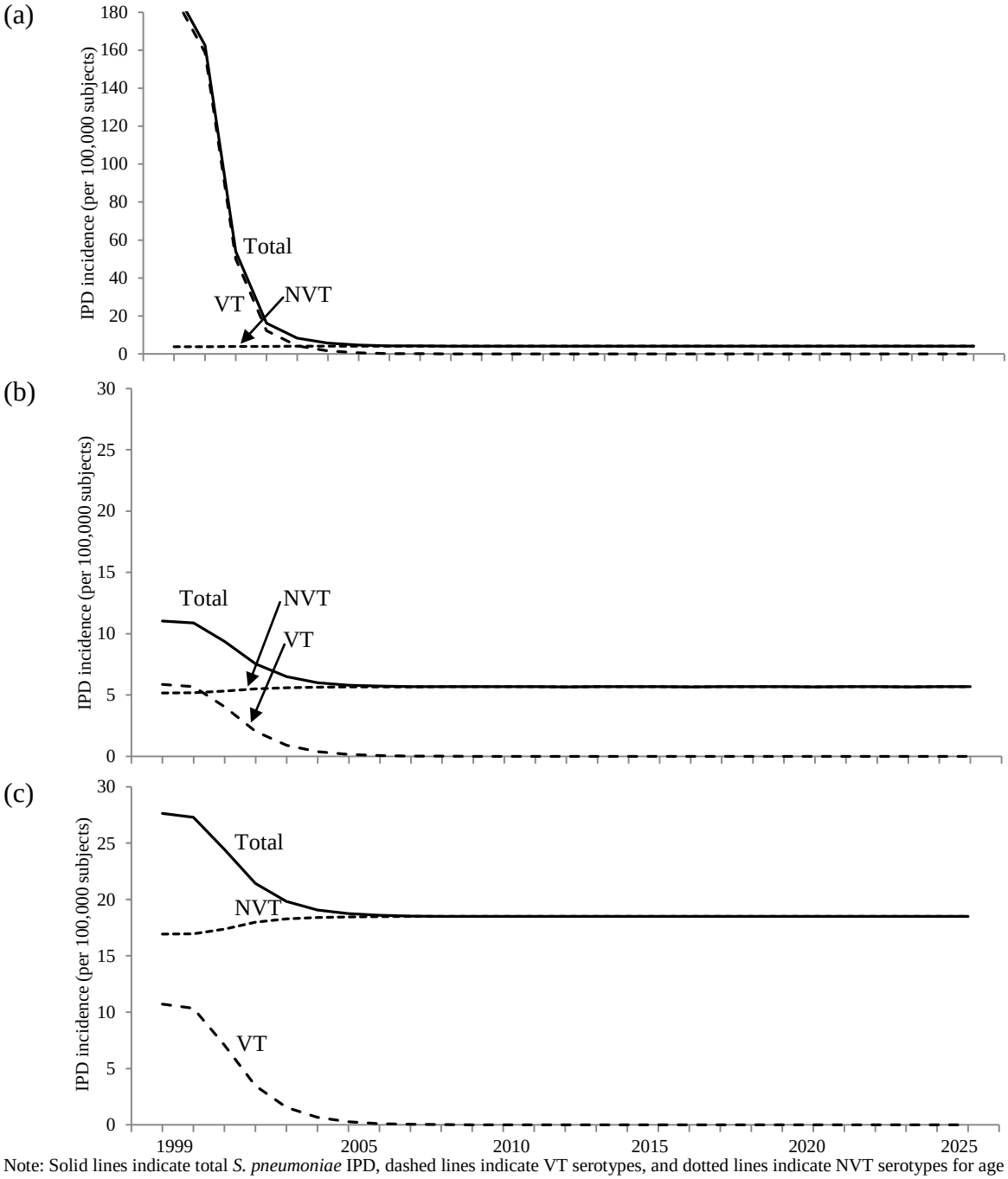
Subject Group	Serotype Group	Colonization Prevalence[a] (%)			IPD Incidence (Cases per 100,000 Subjects)		
		1999	2010	2025	1999	2010	2025
All Subjects	All	12.4	7.2	7.2	28.7	16.3	16.3
	VT	6.4	0.0	0.0	13.8	0.0	0.0
	NVT	6.6	7.2	7.2	14.9	16.3	16.3
<2 yo	All	51.8	33.5	33.5	192.3	4.1	4.1
	VT	29.5	0.0	0.0	188.5	0.0	0.0
	NVT	30.7	33.5	33.5	3.8	4.1	4.1
2 - 15 yo	All	27.5	16.2	16.2	11.0	5.7	5.7
	VT	14.6	0.0	0.0	5.9	0.0	0.0
	NVT	14.8	16.2	16.2	5.2	5.7	5.7
≥16 yo	All	8.7	4.9	4.9	27.6	18.5	18.5
	VT	4.3	0.0	0.0	10.7	0.0	0.0
	NVT	4.5	4.9	4.9	16.9	18.5	18.5

[a] Dually colonized subjects were counted in both VT and NVT serotype groups, allowing the sum of VT and NVT to exceed the total number colonized.



Note: Solid lines indicate total *S. pneumoniae* colonization , dashed lines indicate VT serotypes, and dotted lines indicate NVT serotypes for age groups (a) 0 - < 2 years, (b) 2 - 15 years, and (c) ≥16 years. Scale of vertical axes differs between graphs.

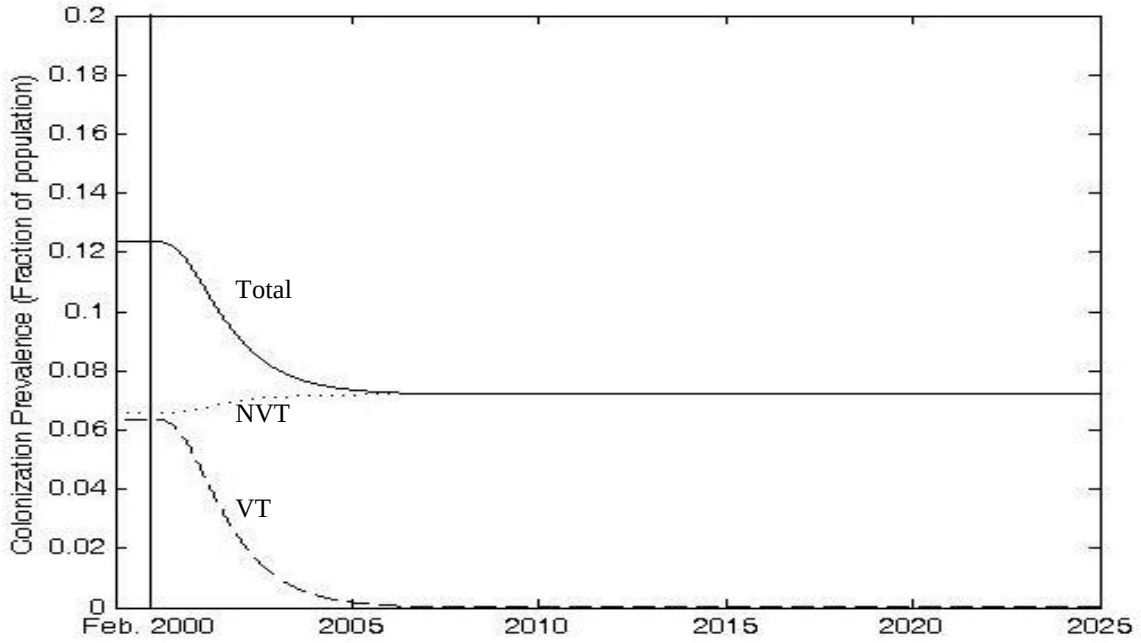
Figure 3.2
Estimated Fractions of Subjects Colonized with *S. pneumoniae*, by Age Group



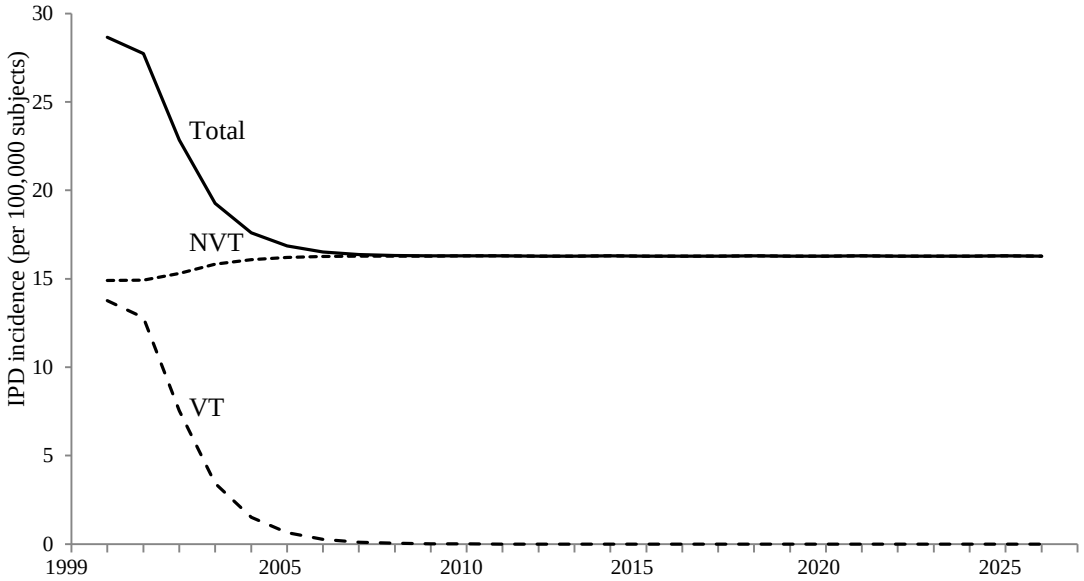
Note: Solid lines indicate total *S. pneumoniae* IPD, dashed lines indicate VT serotypes, and dotted lines indicate NVT serotypes for age groups (a) 0 - < 2 years, (b) 2 - 15 years, and (c) ≥16 years. Scale of vertical axes differs between graphs.

Figure 3.3
Estimated Annual Incidence of IPD per 100,000 Subjects, by Age Group

(a)



(b)



(a) Solid lines indicate total *S. pneumoniae* colonizations, dashed lines indicate VT serotypes, and dotted lines indicate NVT serotypes for entire population.

(b) Solid lines indicate total *S. pneumoniae* IPD, dashed lines indicate VT serotypes, and dotted lines indicate NVT serotypes for entire population.

Figure 3.4
Model-Based Estimates of Total Colonization Prevalence and Total IPD Incidence

groups also reflected slight increases in NVT colonization, with VT colonization dropping below 0.1% of subjects.

The incidence of IPD during the same period also exhibited large decreases. The greatest reduction was in young children, with an overall decrease from 192.3 to 4.1 cases per 100,000 subjects. Reductions in both overall IPD and VT IPD were present in all age groups. Among young children, the percentage of IPD cases caused by VT serotypes decreased from 98% to almost no measurable cases. The percentage of VT-caused IPD cases among older children and adults similarly decreased to less than 0.1 cases per 100,000 subjects. While a portion of the reduction in IPD among older children was due to vaccination, the reduction in IPD among adults and a significant portion of the reduction in IPD among older children reflects the indirect effects of the vaccine. After the virtual eradication of VT colonization and resulting IPD following the introduction of colonization, the estimated fractions of subjects colonized with NVT serotypes, and experiencing NVT IPD, remained approximately constant through 2025.

Sensitivity Analysis

The majority of the parameter values used in this model were obtained from individual studies, and the remaining parameter values were fitted using observed data. While these values represent the best estimates available, these parameter values may or may not be valid for the American population as a whole. Therefore, we conducted a sensitivity analysis to determine which of these parameters had the greatest impact on the model-based estimates. These results could be used to guide future research, in order to provide improved parameter values for use in this model.

The robustness of these results was evaluated using Latin hypercube sampling methodology to generate partial rank correlation coefficients (PRCC) [88]. Latin hypercube sampling allowed us to evaluate the effect of variation in each parameter value with a relatively small

number of samples. This method was the best approach to explore the monotonic, non-linear relationships between parameters and outcomes, to explore the ambiguity surrounding parameter values and to determine which of these parameters had the greatest impact on the outcomes [89].

The two outcomes of interest were the model-estimated prevalence of colonized subjects in 2010, and the model-estimated prevalence of IPD cases per 100,000 subjects in 2010. Relationships between each parameter and these two outcomes were explored graphically. Each parameter in the model, excluding birth, death, and vaccination, was assigned a distribution as described below. Each parameter's distribution was divided into one hundred equiprobable intervals. One hundred complete sets of parameter values were then selected, and each set of values were used in the model to generate the outcomes of interest.

Due to the wide ranges of values used, some combinations of parameter values resulted in models in which one or both of the basic reproductive numbers, R_{01} and R_{02}, were less than 1, which equated to extinction of that serotype group over time. Therefore, the sensitivity analysis was limited to the 68 combinations that had both R_{01} and R_{02} greater than 1, resulting in both serotype groups being present at equilibrium before the introduction of vaccination. The selection of 68 simulations to evaluate the variation in the thirty-eight parameter values exceeded the recommended lower bound of the number of necessary iterations of (4/3 * number of parameters) [90].

The sensitivity analysis was then performed using PRCC to determine which parameters had a significant effect on the outcomes. The parameters which had the most effect on the outcomes were further evaluated to determine the change in outcome caused by a 10% increase in that parameter value.

Distributions of Parameter Values

Distributions for each parameter were selected, as presented in Table 3.5. Parameters whose possible values ranged from 0 to 1, but whose literature-based parameter values were above

0.1 , were assigned a triangular distribution on the [0, 1] interval, with the mode equal to the estimated value. This allowed us to sample a wide range of possible values, with values close to the literature-based parameter value being chosen more frequently.

Parameters whose possible values were on the same range but had literature-based or optimization-based parameter values less than 0.1 (such as the case:colonization fractions [f_{ij}]) were assigned an approximately symmetric Weibull distribution to their inverse. For example, f_{11} had an optimization-based estimate of 0.00017, so 1/ f_{11} was assigned a Weibull distribution centered at 5882. This distribution was truncated so that possible values were less than one. This allowed variation in parameter values without having possible values selected over the entire [0, 1] interval, as would be done with a triangular distribution. This approach was used due to concerns that values close to 1, which were mathematically valid choices but inappropriate biologically, could impact the sensitivity analysis and potentially mask the impact of other variables.

The effect of vaccination on NVT colonization, *k*, was modeled as a beta distribution, to allow a strongly left-skewed distribution with values on the [0, 1] interval. Parameter values for this beta distribution were selected to have 75% of the distribution above 0.75, but also allow some values to be chosen from the lower range. While it is unlikely that *k* has a true value significantly below 1, which would minimize or prevent the post-vaccination increase in NVT colonization and IPD observed in other studies [34; 66; 67], we acknowledge that there is little information available regarding this parameter, so chose a distribution that also included some values over the lower range of mathematically possible values.

All other parameter values were assigned Weibull distributions, with mean values of those distributions equal to the literature-based parameter value. This distribution limited values to be greater than zero, while not constraining the maximum possible value. The shape parameter for all Weibull distributions was set equal to 3.44, to produce an approximately symmetric distribution. In order to explore the effect of different parameter values upon the

outcomes, some parameters whose values were taken to be equal in the baseline parameter set (c_1 and c_2, ρ_1 and ρ_2) were instead generated separately.

Multiple studies have shown the relationship between age and average duration of colonization [9; 10]. Therefore, the values chosen for average duration of colonization were further restricted, within serotype, so that the average duration for adults was less than or equal to the average duration for older children, which was less than or equal to the average duration for young children. All other parameter values were randomly chosen for each set of values used in the sensitivity analysis.

Results of Sensitivity Analysis

The PRCC for parameters and the two outcomes of interest are shown in Table 3.5. All

Table 3.5
Results of the Sensitivity Analysis.

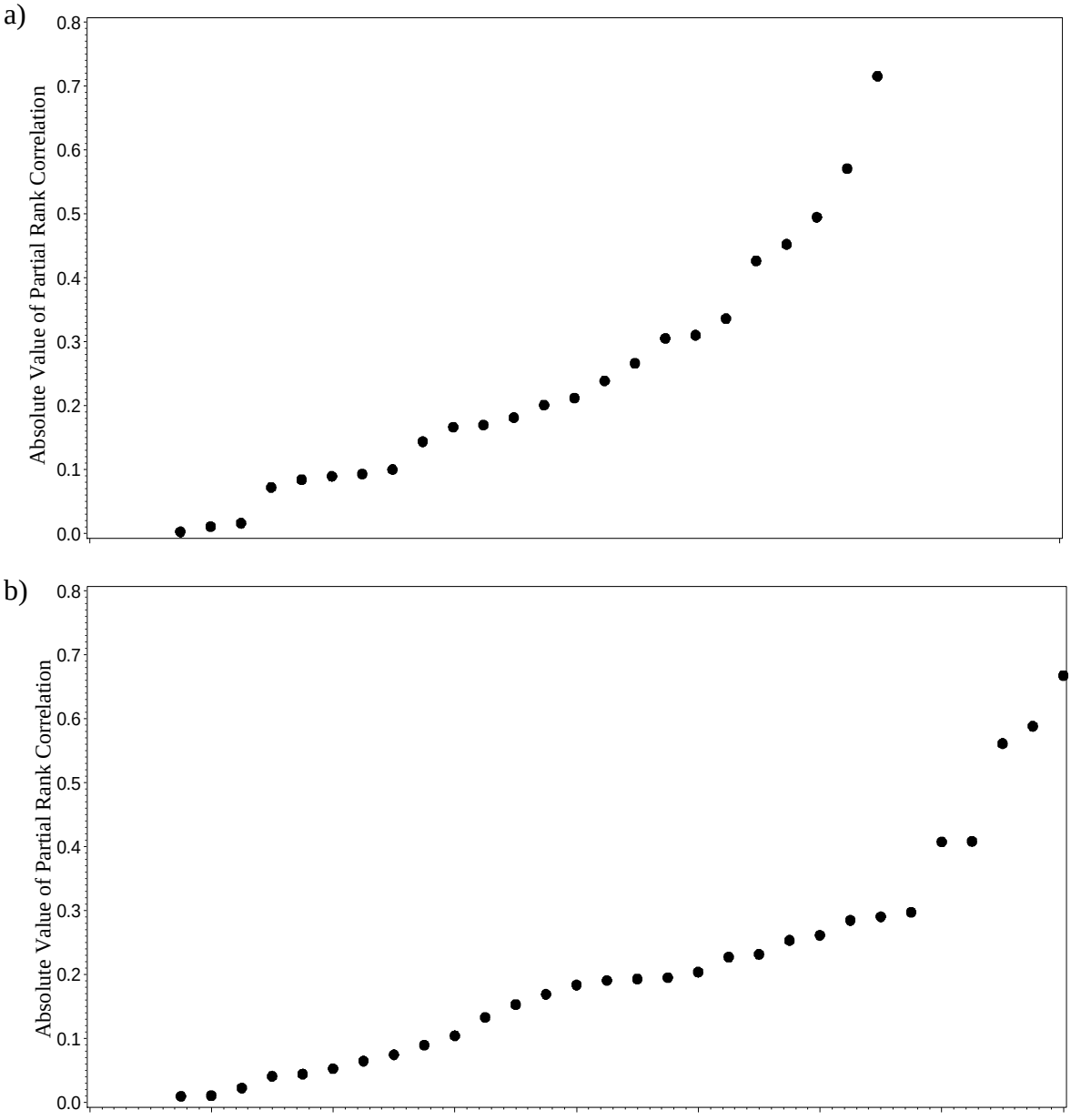
Parameter	Symbol [a]	Distribution [b]	Range of Simulated Values	PRCC Coloni-zation	PRCC IPD
Average weekly contact rates		Weibull			
Among <2 yrs	m_{11}	5.98	1.90-9.49	0.57	0.59
From <2 yrs to 2-15 yrs	m_{12}	20.62	3.63-29.42	N/E	N/E
From <2 yrs to ≥16 yrs	m_{13}	42.55	7.49-60.72	N/E	N/E
Among 2-15 yrs	m_{22}	53.22	15.79-88.15	0.31	0.23
From 2-15 yrs to <2 yrs	m_{21}	2.51	0.83-3.83	0.07	0.01
From 2-15 yrs to ≥16 yrs	m_{23}	49.41	12.22-83.87	0.17	0.19
Among ≥16 yrs	m_{33}	69.14	19.80-105.41	-0.10	-0.02
From ≥16 yrs to <2 yrs	m_{31}	1.16	0.30-1.90	0.18	0.13
From ≥16 yrs to 2-15 yrs	m_{32}	12.21	1.64-24.26	0.14	0.20
Average duration of colonization		Weibull			
Among <2 yrs, VT	γ_{11}^{-1}	17.80	7.2-27.4	0.02	-0.09
Among 2-15 yrs, VT	γ_{12}^{-1}	12.46	3.5-20.3	0.31	0.26
Among ≥16 yrs, VT	γ_{13}^{-1}	8.12	2.5-13.1	-0.21	-0.17
Among <2 yrs, NVT	γ_{21}^{-1}	11.46	4.9-19.2	0.43	0.41
Among 2-15 yrs, NVT	γ_{22}^{-1}	8.01	2.6-12.5	0.08	0. 01
Among ≥16 yrs, NVT	γ_{23}^{-1}	5.23	1.6-8.7	0.17	0.10
Average duration of temporary immunity, VT	ρ_1^{-1}	Weibull 3.79	0.6-6.1	0.27	0.29
Average duration of temporary immunity, VT	ρ_2^{-1}	Weibull 3.79	0.8-6.0	-0.20	-0.18

Table 3.5
Continued

Parameter	Symbol	Distribution [a]	Range of Simulated Values	PRCC Coloni-zation	PRCC IPD
Competition with subject, VT	c_1	Triangular	0.13-0.96	0.09	0.07
Competition with subject, NVT	c_2	Triangular	0.06-0.96	0.09	0.15
Transmission probability, VT		Weibull*			
Among <2 yrs	p_1	49.98	0.013-0.077	0.45	0.41
Among 2-15 yrs and btwn. <2, 2-15 yrs	p_2	875.93	0.00069-0.0051	0.72	0.67
Among ≥16 yrs and btwn. ≥16 yrs, other groups	p_3	3588.5	0.00015-0.0015	0.01	0.19
Ratio of NVT:VT serotype transmission probability	d	Weibull 1.75	0.7-2.8	0.50	0.56
Effect of vaccine on VT colonization	e	Triangular	0.05-0.86	0.34	0.30
Effect of vaccine on NVT colonization	k	Triangular	0.37-1.0	0.24	0.20
Effect of vaccine on VT IPD	r	Triangular	0.01-0.87		-0.04
Fraction of subjects being colonized who develop IPD, VT		Weibull*			
Among <2 yrs	f_{11}	11,494	0.00001-0.0008		0.05
Among 2-15 yrs	f_{12}	111,111	0.00005-0.0003		0.04
Among ≥16 yrs	f_{13}	16,949	0.0002-0.002		-0.06
Fraction of subjects being colonized who develop IPD, NVT		Weibull*			
Among <2 yrs	f_{21}	50,000	0.00001-0.0001		-0.23
Among 2-15 yrs	f_{22}	181,818	0.00003-0.0007		0.29
Among ≥16 yrs	f_{23}	16,949	0.00002-0.003		0.25

[a] Scale parameters were provided for Weibull distributions; mode was provided for triangular distributions. Asterisk indicates Weibull modeled using inverse of parameter value.

relationships were presented in that table, in spite of the number of covariates examined, to allow the reader to review all potentially significant relationships. Previous research has indicated potentially significant covariates based on the *p*-value provided by the analysis [90]. However, the *p*-values decrease as the number of sample sets analyzed increases. Therefore, virtually all covariates could be deemed potentially significant if a large enough number of sample sets was analyzed. We chose a different approach, to allow us to select the most influential parameters while using a cutoff that was not dependent on sample size choice. For each outcome, the absolute values of the partial rank correlation coefficients were sorted from smallest to largest, and graphed as shown in Figure 3.5. The



(a) Relationship between parameters and total colonization after ten years.
(b) Relationship between parameters and total IPD after ten years..

Figure 3.5
Results of Partial Rank Correlation Coefficients, Ranked in Order of Increasing Correlation

five parameters with the greatest impact on each of the outcomes were selected, based on the shape of the graph. In particular, we looked for an obvious break in the graph, in order to select the parameters with the highest impact. The effects of a 10% increase in these parameter values (duration of NVT colonization among young children; transmission probabilities among and between young children and older children; the ratio of transmission probabilities between VT and NVT serotypes, and the average number of contacts among young children) are shown in Table 3.6.

Table 3.6
Effects of Changes in Parameter Values Identified by the Sensitivity Analysis.

Parameter Altered	Parameter Values					2010 Total	
	m_{11}	γ_{21}^{-1}	p_1	p_2	d	Colonization Prevalence [a] (VT, NVT)	IPD Incidence [b] (VT, NVT)
None	5.38	10.3	0.022	0.0013	1.6	1.46% (0.0, 1.46)	0.92 (0.0, 0.92)
m_{11}	5.91	10.3	0.022	0.0013	1.6	1.61% (0.0, 1.61)	1.01 (0.0, 1.01)
γ_{21}^{-1}	5.38	11.3	0.022	0.0013	1.6	1.63% (0.0, 1.63)	1.02 (0.0, 1.02)
p_1	5.38	10.3	0.024	0.0013	1.6	1.61% (0.0, 1.61)	1.01 (0.0, 1.01)
p_2	5.38	10.3	0.022	0.0014	1.6	1.64% (0.0, 1.64)	1.07 (0.0, 1.07)
d	5.38	10.3	0.022	0.0013	1.7	1.83% (0.0, 1.83)	1.27 (0.0, 1.27)

a Dually colonized subjects were counted in both VT and NVT categories.
b IPD incidence reported in number of cases per 100,000 subjects for the entire year.

Discussion

We developed a dynamic transmission model for *S. pneumoniae* that used parameter estimates and reported IPD incidence from published literature to estimate the changes in colonization and IPD in the United States since the PCV7 vaccine was introduced, and to predict future changes. Our model has the advantage of being able to incorporate recently published findings, such as those of Erasto et al. [70], which allowed us to use literature-based estimates of the competition between serotypes, and those of Pilishvili et al. [73], which provided serotype group information for IPD incidence at certain time points. In addition, reported IPD incidence through 2009 was incorporated, allowing this model to be based on data from a longer time period than other recent models.

Model Results

We presented estimates of parameter values for transmission probabilities for each serotype group, and estimated the average duration of temporary immunity. Estimates of these values are not available in the current literature. The sensitivity analysis also provided insight into which parameter values have the greatest impact on model-estimated colonization and IPD levels, and provides direction for future research to refine the most critical parameter values.

Our model explicitly estimated the underlying changes in colonization patterns that have driven the observed changes in IPD levels following the introduction of the vaccine. Our model estimated that VT colonization prevalence would drop to less than 1% of subjects within each age group by 2010, with an increase in NVT colonization to 3-4% higher than pre-vaccination levels within each age group during that time. During the subsequent 16 years, the model predicted that VT colonization would be virtually eradicated, with minimal changes in NVT colonization among any age group. Colonization levels were the driving force behind IPD incidence. Therefore, the model-estimated colonization levels over time were presented to allow further comparisons with other models. Our sensitivity analysis indicated that certain parameters that affected colonization levels were significantly associated with estimated changes in both colonization and IPD, but parameters only affecting IPD (vaccine effect on VT IPD, fractions of colonized subjects developing IPD) had a lesser impact on estimated IPD incidence after controlling for all other variables.

The model also estimated changes in IPD incidence by age group and serotype group. The largest estimated decrease in IPD occurred in young children, with the majority of that decrease occurring within the first five years after the introduction of the vaccine. VT IPD was estimated to become virtually eradicated in young children during the first ten years following the introduction of the vaccine, with an 8% increase in NVT IPD in that age group during the same time. A similar pattern was present among older children, with VT IPD virtually disappearing and NVT IPD increasing by 10%. Due to the indirect effects of the

vaccine, VT IPD was also virtually eradicated among adults. NVT IPD among adults was estimated to increase by 9% by 2010, and to remain virtually unchanged among all age groups between 2010 and 2025.

Model Limitations

Several features of this epidemiological system made it challenging to select the model structure and parameter values. The primary source of data [17], did not provide publicly available information on IPD incidence by VT or NVT serotype group. Most clinical studies of *S. pneumoniae* focus on IPD, as it is the actual cause of the resource utilization in the population, with relatively few studies focusing on the underlying colonization. Several parameter values, such as transmission probabilities and average duration of temporary immunity, are almost impossible to directly estimate in a clinical setting, and have not previously been explicitly estimated using models, although they may have been implicitly estimated as part of a composite parameter. We made assumptions regarding the relationships between transmission probabilities among age groups and serotype groups, in order to reduce eighteen unique parameters down to four unique parameters (p_1, p_2, p_3, d). Similarly, the average duration of temporary immunity was taken to be the same among subjects colonized with all serotypes ($\rho_1 = \rho_2$), in order to reduce the total number of unknown parameter values for the colonization model to five. These five parameter values were then fitted using five known results (fractions of colonized subjects by age group [5], fractions of VT and NVT subjects among colonized young children [81], and time to reinfection in young children [82]). A similar approach was subsequently used to fit the effect of vaccination on VT colonization and fractions of colonized subjects developing IPD (case:colonization ratios), with IPD incidence within age group for each of eleven years and fractions of IPD due to VT serotypes during two of those years being used to fit these seven parameter values (e, f_{ij}). Other parameter values used in the model were based on clinical data in published literature, but were often based on data from a single study, due to the limited amount of available information.

We concede that this approach is not an ideal statistical approach, as above, and that the model estimates would be strengthened if multiple data sources were available for each parameter estimate. In particular, we were overfitting the model by using five data points to fit five unknown values. However, the alternative was to either select values with no numerical basis for one or more parameter estimates, or to impose additional constraints on the parameter relationships, such as assuming $p_1 = p_2$.

In addition, due to the overfitting, we were unable to estimate standard errors for the five fitted parameter values, which does not allow us to determine which estimates were significantly different from zero, or if there were any significant differences between transmission probabilities. However, this type of model is often constrained by the amount of observed data available. Certain parameter values were simply not available, and what information was available was often not provided for specific age groups or serotypes, since much of the data was collected for other purposes and was not geared toward this type of research. This inevitably led to compromises in terms of model structure and parameters. The only other viable alternative, which has been used in similar research, was to propose a number that researchers felt was appropriate, without numeric justification. We felt that, although our estimates may have been based on sparse data, it was preferable to have data-driven estimates instead of estimates without concrete justification.

Another simplification used in this model was the method of grouping multiple serotypes together into the VT and NVT serotype groups, and treating each of these groups as a single serotype, as has been done in similar research [44; 46; 47]. We acknowledge that by grouping serotypes together, all seven VT serotypes were assumed to have the same transmission probabilities and average durations of colonization within age group. The remaining (84+) NVT serotypes were also assumed to have the same transmission probabilities and average durations of colonization, although they differed from the VT serotypes. Very little information was available on specific serotypes; the few studies that

reported information by serotype typically observed only the most common serotypes, and with very few observations of each serotype. Further, incorporating individual serotypes into the model would have vastly increased the number of necessary parameters, with little to no information available to determine those parameter values. We feel that grouping serotypes into VT and NVT groups allowed us to incorporate the changes to the epidemiological system following the introduction of the vaccine, while still having some amount of data available to estimate parameter values.

This model structure also prevented subjects from being colonized with more than one VT serotype, or more than one NVT serotype, concurrently. Such dual colonization has rarely been documented in most populations, possibly due in part to collection mechanism bias [91-93]. The model estimated that 0.7% of subjects were dually colonized with one VT and one NVT serotype prior to vaccination, decreasing below 0.1% within five years. Due to differing colonization levels between age groups, this equated to 8.5% of young children, 1.9% of older children, and 0.2% of adults being dually colonized prior to vaccination, with no measureable fraction of any age group dually colonized by 2010, due to the large decreases in VT colonization during that time.

We acknowledge that some researchers have found subjects dually colonized with more than one VT serotype, or more than one NVT serotype [91; 93]. Our model structure effectively treated a dual VT colonization, or a dual NVT colonization, as a single colonization of that serotype. Therefore, we acknowledge that the model potentially underestimated the average durations of colonization and temporary immunity for these subjects.

We assumed that the effect of vaccination waned by adulthood, providing approximately a 15 year average duration of effect, similar to the waning effects incorporated in other models [44; 46].

The mechanisms behind temporary immunity following colonization are still being investigated. Various research has suggested both serotype-specific and serotype-

independent immunity. Murine-based research found antibody-independent protection lasting at least two months, following colonization or vaccination [74], while an observational study of children suggested serotype-specific protection occurs for some serotypes, but not for others [75]. Earlier studies found that serotype-specific antibodies did not prevent colonization [76]. Although the mechanisms behind immunity are still being explored, we felt it important to include temporarily immune compartments in this model, although some other researchers have chosen to exclude temporary immunity and use SIS models. This decision was supported by the fitted parameter value for the average duration of temporary immunity of 3.4 weeks, which was obtained by optimizing the fit of the model to observed data. If an SIS model were more appropriate than an SIRS model, this fitted parameter value would be expected to be close to zero. However, since we cannot estimate the standard errors for the duration of temporary immunity, we do not know whether that parameter value was significantly different from zero.

Comparisons of Results

Our model predictions differed from those of Temime et al. [44], whose age-structured SIS model incorporated vaccination of 80% of young children. Although there were several differences between these models, such as different pre-vaccination colonization levels, complete prevention of VT colonization by vaccination, constant average durations of colonization across age groups, absence of a temporarily immune phase, incorporation of antibiotic use, and differing infection parameters, the most substantial difference was that the Temime et al. model characterized colonization by one serotype as preventing colonization with the other serotype (effectively, $c_i = 0$), leading to very high serotype replacement. Therefore, when 80% of young children were vaccinated, that model predicted eradication of VT colonization approximately 10 years after the introduction of the vaccine, with NVT colonization levels concurrently increasing by approximately 300%, from about 7% of the population to about 23% of the population (estimated from the figure provided).

More recent research by Melegaro et al. [46] used an SIS model with 100 annual age groups incorporating different average durations of colonization by age group, and then incorporated ratios of IPD cases to numbers of colonized subjects to estimate IPD incidence in the United Kingdom. This model predicted the virtual eradication of VT IPD, and therefore the virtual eradication of VT colonization (due to its characterization of IPD as a fraction of colonization, with no reductions in the chance of developing IPD among subjects who were vaccinated and colonized by VT serotypes), within approximately ten years after the 2006 inclusion of the pneumococcal vaccine into the United Kingdom's routine immunization schedule. It concurrently predicted that the total annual incidence of NVT IPD would increase by approximately 20% within the first ten years, and remain constant after that time. The Melegaro model characterized vaccination as a large (75.6%) reduction in colonization, with no further vaccine-driven reductions in the fraction of colonized subjects developing IPD. It also incorporated larger competition effects than our model, leading to higher serotype replacement, although a more recent publication by many of the same authors indicates that serotype replacement is higher in the United Kingdom than in the United States [71]. In addition, case:colonization ratios were the same for both serotype groups among 1 – 14 year olds, and two to three times higher among NVT serotypes than VT serotypes in the remaining age groups.

In contrast to both the Temime et al. and Melegaro et al. models, our model incorporated a two-part direct vaccination effect, with vaccination causing a 55% reduction in colonization and an additional 89% reduction in the fraction of subjects colonized with VT serotypes who develop VT IPD. This approach incorporated a decrease in colonization consistent with that found in other studies [31; 32; 35; 37] as well as a 95% overall reduction in VT IPD [68; 69]. We found case:colonization ratios to be higher for VT serotypes than NVT serotypes among age group, with the differences between VT and NVT serotypes decreasing in older age groups.

Due to the substantial herd immunity effect, colonization by VT serotypes was predicted to be virtually eradicated, with less than 0.1% of subjects in any age group colonized by those serotypes by 2008. The fractions of subjects colonized with NVT serotypes is estimated to increase by approximately 9% within each age group during the first ten years, with virtually no change in the fractions of subjects colonized after that time.

Although the predicted vast decreases in VT serotype colonization agree with predictions from the Temime et al. and Melegaro et al. models, we predicted smaller increases in NVT serotypes than either of those models. The Temime et al. model predicted approximately a 300% increase in colonization by NVT serotypes within the first ten years following vaccination, with minimal changes in total colonization following vaccination. The Melegaro et al. model estimated approximately a 20% increase in NVT IPD, which indicated a similar increase in NVT colonization, due to its one-part characterization of vaccine effects. Furthermore, that model did not predict IPD as a fraction of the population (such as number of cases per 100,000 subjects), and so did not allow for increases in predicted IPD incidence due to increases in population size (which were likely not necessary for the United Kingdom, which was the focus of that model), making it difficult to predict future prevalence in the United States. Our model predicted NVT IPD incidence to remain approximately constant at 16.3 cases per 100,000 subjects annually from 2009 through 2025. This equates to an increase from 49,000 cases of NVT IPD in 2009 to almost 57,000 cases of NVT IPD in 2025, after incorporating projected population increases as shown in Table B.2 (Appendix B).

A model recently developed by Snedecor et al. [47] used a single-serotype SIS model with 6 age groups focusing on specific types of IPD (invasive pneumonia, bacteremia, or meningitis), and was limited to the first ten years following the introduction of the vaccine. Unlike other models, this model incorporated infection-specific case fatality rates. VT and NVT serotypes were not differentiated, and the 40% reduction in colonization following vaccination, based on the studies of a 9-serotype vaccine among Israeli children [32], was

effectively applied to all serotypes. This model differed significantly from ours in its characterization of the forces of infection, with different age groups' forces of infection having different relationships to the number of infectious subjects in other age groups. Model-based estimates were not presented for colonization, making it difficult to compare results. However, model predictions of overall IPD incidence from the Snedecor et al. model follow the same pattern as those from our model during the ten-year period.

Conclusion

In summary, this model estimated that the largest changes to colonization and IPD would occur during the first ten years following the introduction of the vaccine, with the largest changes observed among young children. Colonization by vaccine-type serotypes were estimated to be virtually eradicated in each age group, with similar decreases occurring in vaccine-type IPD during that time. The prevalence of NVT colonization was estimated to increase within those first ten years, then remain stable during the next 16-year period. The annual incidence of NVT IPD was similarly estimated to increase within the first ten years, and to remain at a constant rate after that time. Total IPD incidence was estimated to decrease from 28.7 cases per 100,000 subjects in 1999 to 16.3 cases per 100,000 subjects in 2005, and to remain constant at that level through 2025.

4. EFFECTS OF ANTIMICROBIAL RESISTANCE ON HOSPITALIZATION COST, DURATION, AND INPATIENT MORTALITY AMONG SUBJECTS WITH *S. PNEUMONIAE* ISOLATES

Streptococcus pneumoniae bacterial infections cause significant medical and economic burdens, both domestically and worldwide. *S. pneumoniae* is a leading cause of community-acquired pneumonia (CAP) [49; 94], as well as being the leading cause of bacterial meningitis [50]. Isolates collected from normally sterile sites, such as the lungs, blood, and cerebrospinal fluid (CSF), are indicative of severe pneumococcal disease.

Influenza and pneumonia were the eighth leading cause of death in the United States in 2007 [95]. In 2008, an estimated 44,000 cases of invasive pneumococcal disease (IPD; pneumococcal disease in the CSF, blood, or other normally sterile site) occurred in the United States, with an estimated 4500 deaths [17]. Non-invasive pneumococcal pneumonia, in which bacteria are not present in the bloodstream, was not included in those estimates of IPD [14]. Therefore, those figures significantly underestimate the total amount of pneumococcal disease.

Cost analyses for pneumococcal pneumonia are often performed using the costs for all pneumonia cases [18]. In 2009, hospitalization costs for subjects with a primary diagnosis of pneumonia were over $9000 per person and 10.5 billion dollars total, among 1.2 million hospitalized patients [96]. Patients with pneumonia have also been shown to present a significant financial burden to employers [51], with the majority of costs attributed to hospitalized patients. While these figures include pneumonia due to other causes, *S. pneumoniae* is the most commonly isolated cause of CAP [49; 94; 97; 98] and is also a significant cause of nosocomial pneumonia.

Several studies have examined mortality or length of stay for some types of *S. pneumoniae* infections, but often include CAP, or are limited to specific age groups, infection sites, or

antimicrobial resistance types. Cost analyses of pneumococcal disease are further complicated by changes over time as different treatments and different medical practices emerge, and as levels of antimicrobial resistance and hospitalization costs increase.

Antimicrobial resistance, often referred to as "drug resistance," occurs when a particular culture of *S. pneumoniae* exhibits decreased susceptibility to a specific group of antimicrobial agents. In general, infections caused by antimicrobial resistant bacteria are believed to cause longer hospitalization times and higher costs when compared with infections by susceptible strains [58]. Resistance can affect patient outcomes by limiting available treatment options, causing a delay in the administration of appropriate therapy, and enhancing virulence. Broad-spectrum antimicrobials, such as fluoroquinolones or third-generation cephalosporins, are now often required for treatment of infections; these agents are typically more expensive, have a greater impact on protective microflora, and can be more toxic or less effective [59]. Furthermore, a patient with a history of resistant bacterial infections may be treated with a very strong antimicrobial, such as vancomycin, when a more narrow-spectrum agent might have worked for that particular infection. This increased use can also contribute to the growth of resistance against these newer, broad-spectrum antimicrobials.

The first known cases of penicillin-resistant *Streptococcus pneumoniae* cultures were identified in the 1960s. Antimicrobial resistance rates have typically been found to be increasing over time [52-56], and vary by geographic area [57]. For instance, in Cleveland, Ohio, the first penicillin-resistant strain was isolated in 1980 and the first macrolide-resistant strain in 1984; by 2004, over 50% of isolates were penicillin-resistant and/or macrolide-resistant and over 25% were lincomycin-resistant [55].

The cost of medical care for specific conditions is a key element of pharmacoeconomics. With so many variables contributing to the medical cost when patients are admitted to the hospital – length of hospitalization , mortality, and antimicrobial resistance especially – it is

often hard to evaluate the expected cost for the "average" patient, which is of vital importance to economic researchers, insurance providers, vaccine manufacturers, and other stakeholders. Accurate estimates of inpatient cost, mortality, and length of stay are vital to the development of economic models to assess the effectiveness of new treatments [60].

There is very little information available about the direct effects of antimicrobial resistance on the costs of pneumococcal disease [18]. While various papers have presented summaries of cost, length of stay (LOS), and/or mortality in patients with *S. pneumoniae* infections or with CAP in general, these figures are typically average values for that sample, leading to estimates that are dependent on the demographic and disease characteristics of those patients. In addition to further exploring the relationships between increasing levels of antimicrobial resistance and cost, LOS, and mortality, we provide estimates of these parameters based on the demographic and disease characteristics shown to be significantly related to each particular outcome. This allows us to provide more accurate estimates of the economic impact of an individual's hospitalization.

METHODS

Data Collection

Medical records reflecting positive *S. pneumoniae* cultures were collected from the Duke University Health System databases for the period of June 1996 through June 2006, and the study approved by the Institutional Review Board. Patients for this study were selected due to the presence of *S. pneumoniae* isolates, regardless of admission reason or primary diagnosis. *S. pneumoniae* isolates were collected from 1153 patients at one or more visits during this period. This population was limited to the isolates that were collected after hospital admission (761 patients) and who had at least one *S. pneumoniae* isolate collected from a location indicative of pneumococcal disease (588 patients). Ten patients who did not have antimicrobial resistance testing conducted were also excluded, resulting in an analysis population of 578 patients. Total cost and/or total charge information was available for 274

of those patients. Information on the 11 patients who had *S. pneumoniae* isolates collected during more than one inpatient hospitalization was limited to the first hospitalization, to ensure independence between records analyzed. Distinct patients were determined by unique medical record numbers.

Source of Isolates

S. pneumoniae isolates were collected from various sites throughout the body. Those isolates collected from locations that would not be expected to have *S. pneumoniae* present during colonization in the absence of pneumococcal disease, including lungs, were classified as being from sterile sites. Isolates collected from the nasopharyngeal area or from areas which could not definitely be classified as being from sterile sites (such as 'abscess' or' tissue') were classified as being from non-sterile sites and excluded from further analyses.

Sterile infection sites were categorized as CSF, blood, lung, or 'other sterile site'. For the purposes of analysis, abdomen, bone, and eye were grouped together as 'other sterile site' locations. Verbatim terms for actual sterile sites are shown in Table 4.1.

Table 4.1
Categorization of Sterile Isolate Locations

Isolate Location Category	Specific Isolate Locations
CSF	Cerebrospinal fluid
Blood	Blood
Lung	Bronchial wash, bronchoalveolar wash, endotracheal suction, tracheal aspirate, pleural
Other Sterile Site	Abdomen: Abdominal, gastric, pancreas, peritoneal Bone: Ankle, finger, hip, knee, neck, sternal, synovial fluid, toe, vertebra, wrist Eye: Aqueous humor, eye, cornea, vitreous fluid

Patients with *S. pneumoniae*-positive isolates collected from more than one location were categorized using a hierarchical method similar to that of Gray [99]. CSF infections are widely considered the most severe; therefore, if a patient had an isolate taken from this location, that patient was categorized as CSF regardless of whatever other isolates were collected. Patients were similarly classified as having infections in the blood, and then

classified as having lung infections, with the remaining patients classified as 'other sterile site'.

Antimicrobial Resistance Testing

*S. pneumonia*e isolates were tested for resistance to different classes of antimicrobials as shown in Table 4.2. All isolates in this analysis population were tested for resistance to at least one, but not more than seven, classes of antimicrobial agents. The majority (445; 77%) were tested for resistance to three or more classes. Ninety-six (17%) and 37 (6%) were tested for resistance to one and two classes of agents, respectively. Isolates were classified as resistant, intermediate, or not resistant (susceptible) to specific agents, with intermediate grouped with resistant for analysis purposes.

Table 4.2
Classes of Antimicrobial Agents

Antimicrobial Class	Antimicrobial Agents
Beta-lactams	Ampicillin, oxacillin, penicillin
Cephalosporins (3rd generation)	Cefotaxime, ceftriaxone
Fluoroquinolones	Gatifloxacin, lefofloxacin, moxifloxacin, ofloxacin
Glycopeptides	Vancomycin
Lincomycins	Clindamycin
Macrolides	Azithromycin, clarithromycin, erythromycin
Sulfonamides	Trimethoprim/sulfamethoxazole

Patients were classified according to dichotomous resistance status (whether or not a patient was resistant to any antimicrobial agent) as well as incremental resistance (resistance to 0, 1, 2, or 3+ classes). Subjects resistant to three or more classes of antimicrobial agents were considered multi-drug resistant (MDR).

Outcomes

Death was defined as inpatient mortality, based on a discharge status of 'Expired' in hospital records, and was unknown for three patients.

The duration of inpatient hospitalization, or length of stay (LOS), was calculated as the discharge date and time minus the admission date and time, divided by 24 hours.

The total costs – i.e., the actual cost of the treatment, supplies, facilities, et cetera – were available for 142 visits. The total charges (amount billed) were available for 191 visits. Both cost and charge were available for 59 of the visits, yielding a total of 274 patients with available monetary information [15].

Analysis Methods

Mortality

Separate bivariate logistic models were created to investigate the individual associations between mortality and categorical age (<18, 18-64, and 65+ years), gender, race (categorized as white, black, and other), isolate location, and dichotomous resistance. Covariates demonstrating significance at $\alpha=0.10$ level in the bivariate model, and all two-way interactions between those covariates, were placed in a multivariate model. A backwards-stepping method was used to remove interactions, then covariates, which were not significant at the $\alpha=0.05$ level, to obtain the final multivariate model.

Length of Stay

LOS and cost outcomes typically have a very skewed distribution, which violates the assumption of normality required for a typical least squares regression, requiring the use of transformations or generalized linear models (GLMs). The following model determination procedure used the method of Manning and Mullahy [100] to determine whether a least squares on the logged outcome or a GLM with a log link provided the best fit to the data.

The multivariate least squares model on the natural logarithm of the outcome, $\ln(LOS)$, was created in the manner previously described, using mortality and patient and disease characteristics as covariates. The model terms for the final multivariate least squares model for the logged outcome were then used to create an equivalent generalized linear model with a log link. Using the Park method described in Manning and Mullahy [100], a subsequent regression of the log of the squared raw-scale residuals on the log of the predicted values was

used to determine whether a Poisson, gamma, or non-linear model was the most appropriate GLM.

The more appropriate model was then determined based on the kurtosis of residuals from each model and the symmetry and variance of residuals from the least squares model. The selected multivariate model was then used to generate estimated values for the outcome for each covariate combination. Mean estimates from least squares models, which are given in the form $E(\ln(y|x))$, were retransformed into $\ln(E(y|x))$ via multiplication by Duan's smearing estimator [101] (the mean of the exponentiated residuals), and then exponentiated to calculate the expected value of the outcome for each combination of the independent variables.

Cost

A least squares regression was used to estimate the relationship between costs and charges among those patients for which both types of information was available, and that ratio used to estimate the actual cost information for patients with only charge information available, and all figures converted to 2010 dollars [102]. The procedure described for LOS was then applied to the 274 subjects with cost information.

Results

Description of Patients

Demographic characteristics of patients are summarized in Table 4.3. Nine percent of subjects had unknown race/ethnicity (n=16) or listed race/ethnicities other than white or black (n=35). Age category and race were correlated (χ^2, p = 0.0001), with white patients being older than black patients and patients of other races.

Table 4.3
Demographics by Antimicrobial Resistance

Characteristic	Value/ Statistic	Number of Resistance Classes					Overall *N*=578
		0 *N*=309	1+ *N*=269	1 *N*=66	2 *N*=77	3+ *N*=126	
Gender, *n* (%)	Female	119 (38.5)	99 (36.8)	25 (37.9)	36 (46.8)	38 (30.2)	218 (37.7)
	Male	190 (61.5)	170 (63.2)	41 (62.1)	41 (53.2)	88 (69.8)	360 (62.3)
Age Category, *n* (%)	< 18 yrs	47 (15.2)	63 (23.4)	9 (13.6)	18 (23.4)	36 (28.6)	110 (19)
	18 – 64 yrs	174 (56.3)	133 (49.4)	36 (54.5)	39 (50.6)	58 (46)	307 (53.1)
	65+ yrs	88 (28.5)	73 (27.1)	21 (31.8)	20 (26)	32 (25.4)	161 (27.9)
Age	*n*	309	269	66	77	126	578
	Mean	48.2	43.8	49.8	44.4	40.2	46.1
	Std Dev	24.14	26.49	23.62	26.8	27.31	25.34
	Median	52	49	52.5	50	46.5	51
	Min/Max	0/96	0/93	0/93	0/89	0/84	0/96
Race/Ethnicity, *n* (%)	White	151 (50.2)	152 (58.2)	36 (54.5)	34 (45.9)	82 (67.8)	303 (52.4)
	Black	134 (44.5)	90 (34.5)	24 (36.4)	36 (48.6)	30 (24.8)	224 (38.8)
	Other*	16 (5.3)	19 (7.3)	6 (9.1)	4 (5.4)	9 (7.4)	35 (6.1)

Note: Asterisk indicated all patients who reported race/ethnicity other than White or Black, excluding Unknown.

Zip code information was available for 562 patients, of which 88% were from North Carolina and 7% from Virginia. Admissions did not significantly vary by year, with the majority of patients (*n*=446, 77%) being admitted via an emergency room.

Thirty-two subjects had positive cultures from the CSF. 265 patients were classified as having *S. pneumoniae* isolates from the blood, and 267 patients as having lung isolates. Relatively few subjects had *S. pneumoniae* isolates only from 'other sterile sites', which were obtained from the abdomen, bone, or eye (4, 4, and 6, respectively).

Isolates from all patients were tested for resistance to beta-lactams, including penicillin; 239 (42%) were resistant. Substantial numbers of subjects yielded isolates which were resistant to third generation cephalosporins (117 of 463 tested, 25.3%), macrolides (119/309, 38.5%), and sulfonamides (138/293, 47.1%). A few isolates were found to be resistant to lincomycin (38/307, 12.4%). No isolates tested for resistance to fluoroquinolones or glycopeptides (256 and 156 patients respectively) displayed any resistance. Additionally, one subject was tested for resistance to tetracycline and was found to be susceptible; due to the very infrequent use of this compound as an antibiotic and the minimal amount of testing for its resistance, this antimicrobial class was excluded from analyses.

Three hundred and nine subjects (53%) provided isolates that were not resistant to any antimicrobial agents against which they were tested. Sixty-six subjects (11%) demonstrated resistance to one agent, and 77 subjects (16% of the 482 patients tested for 2+ agents) showed resistance to two compounds. 126 subjects (28%) of the 445 patients with isolates tested against 3+ agents demonstrated multi-drug resistance (MDR), defined as resistance to three or more antimicrobial agents. Of those, 72 were resistant to 3 compounds, 45 were resistant to four compounds, and 9 were resistant to 5 compounds. Resistance levels, by isolate location, are displayed in Table 4.4.

Table 4.4
Antimicrobial Resistance and Testing, by Isolate Location.

Parameter	#	CSF *N*=32 *n* (%)	Blood *N*=265 *n* (%)	Lung *N*=267 *n* (%)	Other *N*=14 *n* (%)	Overall *N*=578 *n* (%)
Number of Classes of Resistance	0	21 (65.6)	152 (57.4)	126 (47.2)	10 (71.4)	309 (53.5)
	1+	11 (34.4)	113 (42.6)	141 (52.8)	4 (28.6)	269 (46.5)
	1	5 (15.6)	35 (13.2)	26 (9.7)	0	66 (11.4)
	2	5 (15.6)	44 (16.6)	26 (9.7)	2 (14.3)	77 (13.3)
	3+	1 (3.1)	34 (12.8)	89 (33.3)	2 (14.3)	126 (21.8)
Number of Classes Tested	1	3 (9.4)	28 (10.6)	64 (24)	1 (7.1)	96 (16.6)
	2	8 (25)	7 (2.6)	22 (8.2)	0	37 (6.4)
	3	17 (53.1)	82 (30.9)	1 (0.4)	4 (28.6)	104 (18)
	4	1 (3.1)	32 (12.1)	6 (2.2)	0	39 (6.7)
	5	0	16 (6)	72 (27)	5 (35.7)	93 (16.1)
	6	3 (9.4)	98 (37)	98 (36.7)	4 (28.6)	203 (35.1)
	7	0	2 (0.8)	4 (1.5)	0	6 (1)

Mortality

Summary Statistics

One hundred and one patients (18%) with pneumococcal disease died during hospitalization, as shown in Table 4.5. As expected, mortality increased with age, reaching 29.2% among elderly patients. This was slightly more than twice the mortality rate among adults 18-64 years old (14.4%) and three times that of juveniles (9.2%). Deaths were approximately equally distributed by gender, with 18.5% of females and 17% of males expiring during hospitalization. Slightly more whites died than subjects from other races (20.9% among whites; 13.5% among blacks, 11.8% among other races), likely due to the correlation between race and age. Mortality was highest in patients with CSF infections. No obvious trends were present among incremental resistance, with patients without demonstrated antimicrobial resistance having higher mortality than non-susceptible patients (18.2% and 16.9%, respectively).

Table 4.5
Outcome by Antimicrobial Resistance and Isolate Location

Parameter	*N*	Inpatient Mortality *n* (%)	Length of Stay (days) Mean/Median	Cost of Hospitalization* (2010 $) Mean/Median
Overall	578	101 (18)	14.5 / 9.0	35,351 / 17,111
Number of Antimicrobial Resistance Classes				
0	309	56 (18)	12.8 / 7.9	28,109 / 12,671
1+	269	45 (17)	16.4 / 10.3	41,890/ 26,772
1	66	12 (18)	13.4 / 8.1	28,081/ 20,774
2	77	15 (20)	13.5 / 8.8	36,485 / 15,620
3+	126	18 (14)	19.7 / 12.4	50,706 / 32,371
Isolate Location				
CSF	32	8 (25)	10.4 / 9.3	15,864 / 12,163
Blood	265	43 (16)	11.7 / 6.0	24,417 / 9,495
Lung	267	49 (18)	18.1 / 13.8	48,558 / 38,456
Other Sterile Site	14	1 (7)	6.7/ 4.2	18,506 / 7,989

*Cost of hospitalization based on subset with cost or charge information.

Modeling

In bivariate modeling, only categorical age ($p<0.0001$) and race category ($p=0.0606$) were found to be significant at $\alpha = 0.10$. Gender ($p=0.64$), isolate location ($p=0.46$), and dichotomous resistance ($p=0.68$) were not significantly associated with mortality.

The resulting multivariate model included only age, as race and race x age interaction was not significant when age was present in the model. The final model demonstrated appropriate fit, as per the Hosmer and Lemeshow Goodness-of-Fit test ($p=0.9999$).

Patients under 18 years of age were significantly less likely to die during hospitalization than those 65 years or older (odds ratio 0.24, 95% CI 0.12-0.51, $p=0.0002$). Similarly, patients 18-64 years old were also significantly less likely to die during hospitalization than the elderly (odds ratio 0.41, 95% CI 0.26-0.65, $p=0.0002$). The pairwise difference between mortality estimates for juveniles and non-elderly adults was not significant ($p=0.17$). The estimated inpatient mortality rates for patients with pneumococcal disease are: 9.1% (95%

CI, 5.0-16.2) for juveniles, 14.4% (95% CI, 10.9-18.8) for adults, and 29.1% (95% CI, 22.7-36.7) for elderly patients.

Length of Stay

Summary Statistics

Patients with positive *S. pneumoniae* cultures from sterile sites had a mean inpatient hospitalization period of 14.5 days, based on LOS calculations of the total number of hours divided by 24. LOS periods range from 0 to 176.5 days. Mean LOS and median LOS both displayed an increasing trend with increasing amount of antimicrobial resistance, as shown in Table 4.5.

Bivariate Modeling

Isolate location and dichotomous resistance were significant in bivariate modeling at $\alpha = 0.10$, using the logged LOS as the outcome (p-values <0.0001 and .0041 respectively). Gender (p=0.24), race (p=0.26), and age category (p=0.99) were not considered for the multivariate model.

Multivariate Modeling

The multivariate least squares regression model included both isolate location (p<0.0001) and dichotomous resistance (p=0.0310), with their interaction removed due to lack of significance. This model was homoscedastic (p=0.6285). The distribution of the residuals was heavier-tailed than a normal distribution (kurtosis = 2.52, compared to kurtosis=0 for a normal distribution), with variance = 1.09 and skewness of -0.62 (normalized skewness, -6.2).

A generalized linear model (GLM) was then created, using the same model terms with a normal distribution and a log link function. The subsequent regression of the log of the squared residuals on the log of the predicted values produced a coefficient of 0.95, indicating that the Poisson distribution was most appropriate of the GLMs. The GLM using the Poisson distribution and log link was then created.

Per the algorithm given in Manning & Mullahy [100], the least squares model provided the best fit due to variance ≥1 and kurtosis of the residuals. Duan's smearing estimator, calculated as the mean of the exponentiated residuals, was 1.64 [101]. This was multiplied by the predicted values for logged LOS and the product exponentiated, to calculate the estimated LOS as shown in Table 4.6.

Table 4.6
Estimated LOS (days), by Isolate Location and Dichotomous Resistance

Isolate Location	Statistic	Susceptible (0 classes)	Resistant (1+ classes)
CSF	*N*	21	11
	Estimate	11.9	14.4
	95% CI	(7.5-16.3)	(8.9-19.9)
Blood	*N*	152	113
	Estimate	9.7	11.8
	95% CI	(8.3-11.2)	(9.9-13.7)
Lung	*N*	126	141
	Estimate	17.6	21.3
	95% CI	(14.9-20.4)	(18.1-24.5)
Other	*N*	10	4
	Estimate	6.3	7.6
	95% CI	(2.8- 9.7)	(3.3-11.9)

Cost

Summary Statistics

Patients with positive *S. pneumoniae* cultures from sterile sites had inpatient hospitalization costs (converted to 2010 dollars) ranging from $525 to over $311,000. Patients with isolates from the lungs had the highest costs, followed by blood and CSF infections. Patients with resistance had higher average costs than those without any known antimicrobial resistance.

Bivariate Modeling

Gender (p=0.0209), race (p=0.0099), isolate location (p<0.0001), and dichotomous resistance (p=0.0041) showed significant bivariate relationships with logged cost at the α=0.10 level. Age category (p=0.87) was not included in the subsequent models.

Multivariate Modeling

As with LOS, a multivariate least squares regression model was created, using gender, race, isolate location, dichotomous resistance, and all of their two-way interaction terms, using the 271 subjects with all information available. After utilizing the backwards-stepping method, the final model included isolate location (p<0.0001), race (p=0.2206), dichotomous resistance (p=0.0362) and race by resistance interaction (p=0.0011). This final least squares model for the logged cost was not found to be heteroscedastic (p=0.1646). Its residuals were relatively symmetrical (skewness 0.12, normalized skewness 0.78), with variance of 1.07 and kurtosis of 0.39.

Due to the variance (>1) and symmetry (based on low normalized skewness) of the residuals, a least squares regression model is indicated [100]. Duan's smearing estimator was 1.76 [101]. This was multiplied by the predicted values for logged cost and the product exponentiated, to calculate the estimated costs as shown in Table 4.7.

Sensitivity Analysis

We further examined the associations of mortality, LOS, and cost with incremental resistance (0, 1, 2, or 3+ classes). This was considered a sensitivity analysis as not all subjects were tested with 3 or more classes of antimicrobial agents, as shown in Table 4.4.

Mortality

Mortality was not significantly correlated with the incremental resistance (p=0.75).

Table 4.7
Estimated Cost by Race, Isolate Location, and Dichotomous Resistance

Isolate Location	Race	Statistic	Susceptible (0 classes) 2010 $	Resistant (1+ classes) 2010 $
CSF	White	*N*	3	2
		Estimate	17,736	30,811
		95% CI	(4,947- 30,526)	(8,429- 53,194)
	Black	*N*	3	0
		Estimate	24,054	16,875
		95% CI	(6,645- 41,463)	(3,916- 29,834)
	Other	*N*	1	0
		Estimate	18,524	57,311
		95% CI	(0- 37,645)	(752-113,869)
Blood	White	*N*	27	34
		Estimate	15,752	27,365
		95% CI	(11,071- 20,434)	(19,907- 34,823)
	Black	*N*	38	23
		Estimate	21,363	14,987
		95% CI	(15,269- 27,458)	(10,062- 19,913)
	Other	*N*	1	3
		Estimate	16,452	50,900
		95% CI	(2,013- 30,891)	(14,827- 86,973)
Lung	White	*N*	32	50
		Estimate	41,632	72,322
		95% CI	(29,626- 53,638)	(54,425- 90,219)
	Black	*N*	15	23
		Estimate	56,462	39,610
		95% CI	(37,796- 75,128)	(26,584- 52,637)
	Other	*N*	3	6
		Estimate	43,482	134,523
		95% CI	(5,866- 81,098)	(41,439-227,608)
Other Sterile	White	*N*	2	0
		Estimate	15,723	27,314
		95% CI	(2,854- 28,593)	(4,558- 50,071)
	Black	*N*	3	1
		Estimate	21,324	14,960
		95% CI	(4,102- 38,546)	(2,374- 27,545)
	Other	*N*	1	0
		Estimate	16,422	50,806
		95% CI	(0- 34,127)	(0-104,547)

LOS

Incremental resistance was significantly associated with logged LOS in the bivariate model ($p<0.0001$). The subsequent multivariate least squares model included both isolate location ($p<0.0001$) and incremental resistance ($p=0.0251$), was homoscedastic ($p=0.5015$), skewed (skewness -0.61; normalized skewness -5.95), and had variance of 1.08. The associated generalized linear model was created, and the subsequent regression using the Park method resulted in a coefficient of 1.29, indicating a Poisson model as before. The log-scale errors were heavy-tailed (kurtosis 2.4). As with the dichotomous resistance model, the least squares model was concluded to be the best fit per the selection algorithm [100]. Estimated LOS values from this model were given in Table 4.8. Due to limited information available for some combinations, some confidence intervals were very wide. Predictions for the susceptible class were slightly different from those in Table 4.6, due to model differences.

Table 4.8
Estimated LOS (days), by Isolate Location and Dichotomous Resistance

Isolate Location	Statistic	Susceptible (0 classes)	Number of Resistance Classes		
			1	2	3+
CSF	*n*	21	5	5	1
	Estimate	12.3	13.5	12.8	17.4
	95% CI	(7.8-16.8)	(7.7-19.3)	(7.4-18.1)	(10.1-24.7)
Blood	*n*	152	35	44	34
	Estimate	9.9	10.9	10.3	14
	95% CI	(8.4-11.4)	(8.0-13.8)	(7.7-12.8)	(10.8-17.2)
Lung	*n*	126	26	26	89
	Estimate	17.1	18.8	17.8	24.2
	95% CI	(14.4-19.8)	(13.6-24.0)	(13.1-22.4)	(19.6-28.8)
Other	*n*	10	0	2	2
	Estimate	6.3	6.9	6.5	8.9
	95% CI	(2.8- 9.7)	(2.7-11.1)	(2.7-10.3)	(3.7-14.0)

Cost

Incremental resistance was significantly associated with logged cost in the bivariate model (p=0.0013). The subsequent multivariate least squares model included isolate location (p<0.0001), race (p=0.0031), incremental resistance (p=0.0202), and race x resistance class interaction (p = .0236) was homoscedastic (p=0.2735), relatively unskewed (skewness 0.11; normalized skewness 0.72), and had variance of 1.04. The associated generalized linear model was created, and the subsequent regression using the Park method resulted in a coefficient of 1.24, indicating a Poisson model as before. The log-scale errors from the Poisson model had approximately normal tails (kurtosis = -0.04). As with the dichotomous resistance model, the least squares model was concluded to be the best fit per the selection algorithm [100]. Estimated LOS values from this model are given in Table 4.9. Predictions for the susceptible class were slightly different from those in Table 4.6, due to model differences.

Discussion

It is well known that mortality rates increase with age; therefore, it was anticipated that categorical age would be related to mortality. Categorical age and race were correlated in this sample, with a large portion of elderly patients being white. Therefore, the bivariate relationship between race and mortality was confounded by the correlation between age and race, and the race/mortality relationship was better summarized by controlling for age through the multivariate model.

Although isolate location category was not significantly associated with mortality in our data, the data reflected the highest mortality (8/32, 25%) among subjects with CSF isolates, which is similar to CSF mortality rates found in other studies [103; 104]. The mortality rate for patients with bloodstream infections (43/265, 16%) was well within the rates of 11-32% [105-113] found in other research, and the mortality rate for lung infections (49/267, 18%)

Table 4.9
Estimated Cost, by Race, Isolate Location, and Incremental Resistance

Location		Susceptible	Number of Resistance Classes		
Race	Statistic	(0 classes)	1	2	3+
CSF					
White	*n*	3	1	1	0
	Estimate	18,236	26,220	23,835	37,254
	95% CI	(5,080- 31,391)	(4,709- 47,730)	(3,834- 43,837)	(8,726- 65,782)
Black	*n*	3	0	0	0
	Estimate	24,642	10,826	22,606	21,015
	95% CI	(6,811- 42,473)	(1,249- 20,403)	(2,348- 42,864)	(3,000- 39,030)
Other	*n*	1	0	0	0
	Estimate	19,024	46,936	74,059	59,316
	95% CI	(0- 38,626)	(0-122,723)	(0-194,110)	(0-128,107)
Blood					
White	*n*	27	7	9	18
	Estimate	15,455	22,221	20,200	31,573
	95% CI	(10,865- 20,045)	(11,130- 33,312)	(9,571- 30,829)	(20,980- 42,166)
Black	*n*	38	7	9	7
	Estimate	20,884	9,175	19,158	17,810
	95% CI	(14,940- 26,828)	(4,140- 14,210)	(8,492- 29,825)	(8,726- 26,894)
Other	*n*	1	1	0	2
	Estimate	16,122	39,778	62,765	50,270
	95% CI	(2,011- 30,234)	(0- 97,745)	(0-155,379)	(3,433- 97,106)
Lung					
White	*n*	32	11	6	33
	Estimate	40,174	57,762	52,510	82,072
	95% CI	(28,597- 51,750)	(29,802- 85,722)	(24,202- 80,817)	(57,246-106,897)
Black	*n*	15	8	5	10
	Estimate	54,287	23,850	49,801	46,297
	95% CI	(36,324- 72,250)	(10,866- 36,834)	(21,175- 78,428)	(23,217- 69,377)
Other	*n*	3	1	2	3
	Estimate	41,909	103,401	163,153	130,673
	95% CI	(5,759- 78,059)	(0-254,082)	(0-399,904)	(9,934-251,412)
Other					
White	*n*	2	0	0	0
	Estimate	14,929	21,465	19,513	30,499
	95% CI	(2,722- 27,135)	(1,502- 41,427)	(949- 38,078)	(4,462- 56,535)
Black	*n*	3	0	0	1
	Estimate	20,173	8,863	18,507	17,204
	95% CI	(3,893- 36,454)	(340- 17,386)	(514- 36,499)	(1,728- 32,680)
Other	*n*	1	0	0	0
	Estimate	15,574	38,425	60,629	48,559
	95% CI	(0- 32,324)	(0-102,139)	(0-161,543)	(0-107,782)

was comparable to the pneumococcal pneumonia mortality rates of 8-26% [94; 111; 114-116], although reported pneumococcal pneumonia mortality rates typically include subjects with invasive pneumonia (i.e. bacteria were found in both the blood and lung).

Estimated LOS for patients displaying antimicrobial resistance was consistently longer than that for susceptible patients, with the longest LOS for patients exhibiting MDR, regardless of isolate location. LOS also varied by isolate location, with estimated LOS among subjects with lung infections significantly higher than those with blood ($p<0.0001$) infections, after controlling for resistance. Estimated LOS among patients with lung infections was significantly longer than that for patients with CSF infections when controlling for dichotomous resistance ($p=0.0466$) but not when incorporating incremental resistance ($p=0.0961$).

After controlling for isolate location, white patients and "other" patients had higher estimated hospitalization costs in the presence of antimicrobial resistance, while black patients with antimicrobial resistance had lower estimated costs than susceptible black patients. Further research is needed to fully understand this pattern, which may be present due to differences in comorbidities.

It should be noted that most research of this type is based on the use of hospitalization costs and a cost-to-charge ratio conversion [117; 118] such as the Medicare average cost-to-charge ratio of 0.57 [119]. By contrast, this method of using actual charges better reflects the true cost of hospitalization. While charges may provide a proxy for costs, they generally overestimate costs, and the relationship between charges and costs varies between hospital departments [60].

As shown in Table 4.4, isolate testing for antimicrobial resistance varied between patients. Therefore, care should be taken when interpreting the results based on incremental resistance, since these results may be confounded by other factors.

Mortality was found to be associated with LOS, but was not associated with cost, in this sample. However, it was not included as a covariate in LOS or cost models, as the intent is to provide estimates based on characteristics at admission.

This study only evaluated outcomes from the first hospitalization during the data collection period in which a *S. pneumoniae* isolate was collected from a sterile site in a patient. It is possible that some deaths, which occurred after discharge, were not included. However, other research has found that more than half of the fatalities were during the first seven days [106; 120], and more than two-thirds of deaths were within the first 10 days after admission [106]. In addition, it seemed unlikely that a patient with severely declining health would be discharged, so we anticipate that the vast majority of deaths due to *S. pneumoniae* would be captured using inpatient mortality. Mortalities were also affected by patients having do-not-resuscitate orders, which are more common among the elderly or those with severe comorbidities.

This study only considered medical costs (both direct and indirect inpatient costs) from patients who were admitted to the hospital, and excluded outpatient costs, repeat hospitalizations, prescription drugs, physician office visits, and other economic factors such as loss of work and disability costs for patients who were hospitalized. One study found that hospitalizations comprise 87-95% of the total medical costs [18]. Another study estimated inpatient costs for all-cause pneumonia to be approximately 63% of health care costs and 55% of total costs (including health care, disability, and absenteeism costs) [121]. In addition, estimates of inpatient treatment range from one in three to one in five patients for pneumonia [18; 122; 123]. It should be emphasized that these cost estimates reflected only those patients who were ill enough to be admitted to the hospital, and did not include the sometimes substantial additional costs for outpatient care and other costs. The mortality rate for CAP patients who do not require hospitalization is less than 1% [94; 124], and elderly patients with pneumococcal CAP have been found to have significantly lower mortality rates

than those with CAP due to other causes [125]. Deaths among those who were not hospitalized were therefore considered negligible.

Additionally, the inpatient population may not have been representative of all patients with pneumococcal disease; patients with CAP who were treated on an outpatient basis tend to be younger, have fewer comorbidities, and have less severe infections than those who were treated on an inpatient basis [94].

This data was collected from a single health system, and may not be representative of other health systems or geographic regions.

Conclusion

No significant relationship was found between *Streptococcus pneumoniae* disease characteristics and mortality, which was significantly associated with age. The longest estimated LOS and highest estimated costs were associated with patients with lung infections. Estimated LOS was consistently higher among resistant patients, as were estimated costs among white patients and those of 'other' races. Estimated costs among black patients were higher among susceptible patients. Insufficient information was available to determine the cause of the pattern difference between racial groups, which was possibly due to underlying differences in health status between the different groups. Patients with multi-drug resistance consistently had the highest estimated costs.

5. COST-BENEFIT ANALYSIS OF THE SEVEN-SEROTYPE PNEUMOCOCCAL CONJUGATE VACCINE, UTILIZING ESTIMATES FROM A DYNAMIC TRANSMISSION MODEL

Introduction

Multiple cost-effectiveness analyses have been performed on the PCV7 pneumococcal conjugate vaccine since its February 2000 approval in the United States. Indirect vaccination effects (herd immunity) were typically underrepresented in these analyses, due to lack of available information or use of model types that did not incorporate these effects. Many of the analyses conducted were cost-effectiveness or cost-utility analyses, which estimate effectiveness in terms of cost per life-years saved [62; 64] or cost per quality-adjusted life-years saved [63], and then require subsequent evaluation to determine whether cutoff criteria were met (i.e., whether the intervention is a good use of resources). In contrast, cost-benefit analyses, such as that used in this research, evaluate the consequences of a vaccination program from a purely monetary standpoint, and identify whether the monetary benefits outweigh the costs.

We developed an age-structured dynamic transmission model to explore the effect of vaccination on pneumococcal colonization and invasive pneumococcal disease (IPD) in the United States (Chapter 3). Reported cases of IPD, as tracked by the Centers for Disease Control and Prevention's Active Bacterial Core surveillance (CDC ABCs) team, consists primarily of meningitis, bacteremia, and invasive pneumonia (i.e., pneumonia in which bacteria is also present in the blood). Non-invasive pneumococcal pneumonia (i.e., pneumonia in which bacteria has not spread into the bloodstream) is not classified as IPD by the CDC [14]. The transmission model was calibrated to CDC-reported IPD, so model-based estimates of IPD also do not include non-invasive pneumonia.

Prior to the introduction of the PCV7 vaccine, the CDC estimated that *Streptococcus pneumoniae* caused 3000 cases of meningitis, 50,000 cases of bacteremia, and 500,000 cases

of pneumonia annually [19; 126]. Therefore, although significantly underrepresented in estimates of pneumococcal disease based on CDC-reported IPD, pneumococcal pneumonia constitutes the majority of pneumococcal disease. Consequently, reductions in pneumococcal pneumonia are a driving force in reductions in medical costs after the introduction of pneumococcal vaccination.

Fractions of model-based estimates of reported IPD were used to estimate the incidence of pneumococcal meningitis and bacteremia, by age group and year. Due to the underrepresentation of pneumonia in model-based estimates of IPD, model-based estimates of colonization were used to develop estimates of pneumococcal pneumonia incidence, by age group and year.

Hospitalization costs (Chapter 4) and vaccination costs were combined with model-based estimates of pneumococcal disease to perform a cost-benefit analysis, using the following approach. The total number of doses of vaccine administered through 2009 was estimated, and the total cost of vaccination calculated. The incidence of pneumococcal disease in 1999 (pre-vaccination) and during the first ten years of PCV7 vaccination (2000 – 2009), by disease type and age group, was then calculated. Fractions of model-based estimates of IPD were used to calculate the incidence of pneumococcal meningitis and bacteremia, and pneumococcal pneumonia incidence was calculated as a fraction of the model-based estimate of colonized subjects. The amount of pneumococcal disease prevented was calculated as the difference between the total number of pneumococcal disease cases during that ten-year period and the estimated number of cases that would have occurred during that time in a hypothetical no-vaccine scenario. The average cost of each pneumococcal disease case was estimated, and the total cost savings associated with vaccine use calculated. The future cost savings for continued use of this vaccine through 2025 was also predicted. A supplemental analysis was also conducted using hospitalization costs based on a nationwide inpatient sample.

Vaccination and Population

The total population of the United States by age group (young children, <2 years; older children, 2-15 years; adults, ≥16 years) was calculated based on information available from CDC National Vital Statistics System (NVSS) [127] and the U.S. Census Bureau Population Division (CBPD) [128]. Population estimates for all three age groups for 1999 were taken from CBPD data. Estimates of the numbers of young children and older children for 2000 - 2009 were calculated as the average of the NVSS and CBPD estimates. Adult population sizes were taken directly from CBPD data, which includes a specific estimate for ≥16 year olds. These annual population estimates are displayed in Appendix B.

The CDC recommends that children receive 4 doses of pneumococcal conjugate vaccine, at 2, 4, 6, and 12-15 months of age [129]. The number of children eligible to receive vaccination was taken to be the number of children alive at 1 year of age. This was calculated as the number of live births minus the number of deaths of infants less than 1 year old, for each year (Appendix B).

The percentage of eligible children receiving 1, 2, 3, or 4 doses of pneumococcal vaccine was estimated from the National Immunization Survey for 2001 - 2009 [84]. Where available, vaccination information based on actual birth cohort was used [85]. The number of doses received by children in the remaining birth cohorts was calculated using survey sampling methodology provided by the CDC [130]. Weighted averages of the results for each survey year were used to estimate vaccination for each birth cohort. Estimates of percentages of children vaccinated and numbers of doses administered are presented by year in Table 5.1. Due to the age at which the survey is administered, complete vaccination information was not available after the 2006 birth cohort. Therefore, vaccination levels were taken to remain constant at 2006 levels in later birth cohorts. Overall, almost 139 million doses were estimated to have been administered to children born through 2009.

The cost of the vaccine was obtained from the CDC price list [61], using the cost per dose for the private sector each year, and converted to 2010 dollars [102]. The cost per dose for vaccine administered to each birth cohort was conservatively taken to be equal to the cost listed for the subsequent calendar year. In other words, children born in 1999 were assigned the vaccine cost for 2000, and those born in 2009 were assigned the vaccine cost for 2010. Children born in 1998 were also assigned the vaccine cost for 2000. Twelve dollars per dose for vaccine administration costs, based on estimated costs in Ray et al. [64], were also included in the cost per dose presented in Table 5.1.

Table 5.1
Estimates of Eligible Children and Vaccination

		% Eligible Children Receiving Vaccine						
Year	Number Eligible for Vaccination	1 Dose (%)	2 Doses (%)	3 Doses (%)	4 Doses (%)	Total Doses	Cost Per Dose [b]	Total Cost (Millions of Dollars)
1998	3,913,174	6.2	4.4	0.5	0.1	661,000	$85.45	$57
1999	4,130,275	14.4	19.2	8.0	4.9	3,980,000	$85.45	$340
2000	4,030,779	9.7	16.4	23.7	26.2	8,810,000	$84.34	$743
2001	3,998,365	5.9	12.5	30.2	38.0	10,940,000	$83.21	$910
2002	3,993,692	4.5	9.4	27.5	48.5	11,970,000	$85.06	$1018
2003	4,061,925	3.0	6.3	25.0	58.7	13,220,000	$83.17	$1100
2004	4,084,116	2.6	4.3	15.7	71.8	14,120,000	$85.63	$1209
2005	4,109,909	1.8	3.0	12.5	78.5	14,770,000	$91.72	$1355
2006	4,237,028	1.4	2.6	11.4	80.6	15,380,000	$95.28	$1465
2007 [a]	4,287,095	1.4	2.6	11.4	80.6	15,560,000	$96.95	$1509
2008 [a]	4,219,665	1.4	2.6	11.4	80.6	15,320,000	$97.26	$1490
2009 [a]	4,104,493	1.4	2.6	11.4	80.6	14,900,000	$120.75[c]	$1799
Total	...	...	...	...	...	139,631,000	...	$12,995

a Estimates after 2006 were assumed to be constant at 2006 levels.
b Costs converted to 2010 dollars and presented for the subsequent calendar year. Costs included the cost of vaccine plus $12 administration cost.
c Vaccine cost associated with the PCV13 vaccine, introduced in February 2010, which includes six serotypes in addition to the seven included in the PCV7 vaccine.

During the first ten years of vaccine administration, an estimated 13.0 billion dollars were spent on the vaccine and its administration. Our approach included vaccination costs for 11 birth cohorts in the summary of vaccine costs, for comparison with cost savings over a ten-year period. It also assigned the price associated with the new PCV13 vaccine (a 13-serotype

vaccine which includes the serotypes found in the PCV7 vaccine), introduced in February 2010, to children in the 2009 birth cohort. Therefore, these figures provide a slight overestimate of actual vaccine costs during this ten-year period.

Pneumococcal Disease

The dynamic transmission model described in Chapter 3 was used to estimate both colonization and IPD for 1999 (pre-vaccination) and the first ten years of vaccine use in the United States (2000 – 2009), as shown in Tables C.1 and C.2. Model-based estimates of IPD were calibrated to IPD incidence reported by the CDC ABCs. This reported IPD included meningitis (isolates collected from the cerebrospinal fluid [CSF]), bacteremia (isolates collected from the blood), invasive pneumonia (isolates collected from both the blood and lungs), and other types of disease. Non-invasive pneumonia, in which bacteria were not isolated from the blood, was not included in these reports of IPD incidence.

Meningitis and bacteremia (excluding subjects with invasive pneumonia) accounted for approximately 25% of reported IPD. Invasive pneumococcal pneumonia accounted for approximately two-thirds of reported IPD. Since only 10-25% of pneumonia spreads to the bloodstream [126], the majority of pneumonia cases were not included in CDC-reported IPD. Furthermore, many cases of pneumonia are treated without the underlying microbiological cause being determined, causing pneumococcal pneumonia to be under-diagnosed. Therefore, pneumonia incidence could not be accurately calculated from estimates of reported IPD, and was instead calculated as a fraction of model-estimated colonization (Table C.2).

Pneumococcal Meningitis

The number of meningitis cases reported to the CDC ABCs, by age group and two-year period (1998-99, 2000-01, 2002-03, 2004-05) was presented by Hsu et al. [27]. The fraction of reported IPD caused by meningitis, within age group, was then calculated using these breakdowns by age group in conjunction with model-based estimates (Chapter 3) of reported

IPD cases [17]. Percentages of reported IPD caused by meningitis ranged from 5.1% to 6.3% for young children, from 4.7% to 7.9% for older children, and from 4.0% to 5.2% for adults. An analysis of variance model did not find that the changes in percentages between different two-year periods were statistically significant. Therefore, mean values of 5.35%, 6.22%, and 4.38% were used to calculate the fraction of IPD caused by meningitis within each age group. The estimated numbers of meningitis cases by age group, as well as the numbers of cases per 100,000 subjects, are presented in Table 5.2.

Table 5.2
Model-Estimated Meningitis and Bacteremia Incidence, by Age Group and Year

	Meningitis				Bacteremia			
	Rate, by Age Group [a]				Rate, by Age Group [a]			
Year	<2 years	2-15 years	≥16 years	Overall Number of Cases (Rate [a])	<2 years	2-15 years	≥16 years	Overall Number of Cases (Rate [a])
1999	10.3	0.7	1.2	3,700 (1.4)	40.8	2.3	5.9	16,700 (6.1)
2000	8.7	0.7	1.2	3,700 (1.3)	34.5	2.3	5.8	16,800 (5.9)
2001	2.9	0.6	1.1	2,930 (1.0)	11.4	2	5.2	13,500 (4.7)
2002	0.9	0.5	0.9	2,430 (0.8)	3.4	1.6	4.5	11,300 (3.9)
2003	0.4	0.4	0.9	2,220 (0.8)	1.7	1.4	4.2	10,500 (3.6)
2004	0.3	0.4	0.8	2,140 (0.7)	1.2	1.3	4	10,000 (3.4)
2005	0.2	0.4	0.8	2,120 (0.7)	1	1.2	4	10,000 (3.4)
2006	0.2	0.4	0.8	2,120 (0.7)	0.9	1.2	3.9	10,000 (3.3)
2007	0.2	0.4	0.8	2,130 (0.7)	0.9	1.2	3.9	10,100 (3.3)
2008	0.2	0.4	0.8	2,150 (0.7)	0.9	1.2	3.9	10,200 (3.3)
2009	0.2	0.4	0.8	2,170 (0.7)	0.9	1.2	3.9	10,300 (3.4)

a Rates indicated number of cases per 100,000 subjects, within each age group.

Pneumococcal Bacteremia

While normally considered a less severe illness than pneumococcal meningitis, pneumococcal bacteremia is much more prevalent. We considered subjects who were categorized by the CDC as having bacteremia without focus to have bacteremia; those categorized as having pneumonia with bacteremia were considered to have invasive pneumococcal pneumonia, and were not incorporated into our estimates of bacteremia.

The fraction of reported IPD cases caused by bacteremia was significantly higher in 1999 (37.6%) and 2000 (30.3%) than in 2001 – 2009, when it ranged from 19.2% to 23.0%. This

was likely due to higher levels of reporting prior to the introduction of the vaccine, similar to that observed in the United Kingdom [71]. Therefore, we used the observed fractions of 37.6% and 30.3% of reported IPD for 1999 and 2000, respectively, in conjunction with model-estimated IPD incidence (Table C.2) to estimate the incidence of pneumococcal bacteremia within age group for those two years. Bacteremia incidence in subsequent years was similarly calculated, using the 2001 – 2009 average of 21.4% of reported IPD being caused by bacteremia. The estimated incidence of bacteremia, by year and age group, is shown in Table 5.2.

Pneumococcal Pneumonia

The majority of pneumococcal pneumonia is not invasive [15], so is not included in CDC-reported IPD. Model-based estimates of IPD were calibrated to reported IPD levels, and similarly underrepresented the incidence of pneumococcal pneumonia. Furthermore, a clinical diagnosis of pneumonia does not require microbiologic testing to determine the underlying cause. Therefore, many, if not most, cases of pneumonia are treated without the etiology being determined, as confirmed by subject matter experts in the CDC's Division of Infectious Diseases. Pneumococcal pneumonia is then under-diagnosed, making it problematic to estimate its incidence using specific diagnoses of pneumococcal pneumonia.

Therefore, its incidence was estimated using the following method. The 1999 Healthcare Cost and Utilization Project (HCUP) survey identified 2,221,001 hospital discharges with pneumonia listed as one of the diagnoses, as defined by a Clinical Classification Software (CCS) code of 122 [96]. Each hospital discharge was taken to be a unique subject. Results from the six different age groups provided were combined to calculate that 92,014 children less than 2 years of age, 153,400 children 2 – 15 years of age, and 1,973,935 adults at least 16 years of age were hospitalized for pneumonia. The 1651 records with unknown age were not included in these calculations.

The HCUP records included both community-acquired pneumonia (CAP) and nosocomial, or hospital-acquired, pneumonia. Nosocomial pneumonia has been estimated to occur in 4 – 8 subjects per 1000 hospitalizations [131]. Therefore, 0.006% of hospitalized subjects with a diagnosis of pneumonia were taken to have nosocomial pneumonia, which was calculated to be 552, 920, and 11,844 cases, from youngest to oldest age group.

Sixty-two percent of subjects with CAP were taken to have been admitted to the hospital during their illness, as calculated from a retrospective study of 2287 patients with CAP [94]. The total number of cases of CAP, including both hospitalized and outpatient subjects, was calculated by dividing the CAP incidence by 0.62.

The fraction of CAP due to *S. pneumoniae* was calculated as follows. The total number of subjects with a diagnosis of unspecified bacterial pneumonia, unspecified bronchopneumonia, or unspecified pneumonia (ICD-9-CM codes of 482.9, 485, and 486, respectively) was subtracted from total number of subjects with a pneumonia diagnosis, to obtain the total number of patients with pneumonia due to a known cause. This number included patients with diagnoses of pneumonia with partially specified etiology (e.g., unspecified viral pneumonia) that ruled out *S. pneumoniae* as the cause. The total number of patients diagnosed with pneumococcal pneumonia (ICD-9-CM: 481) was divided by the number of patients diagnosed with pneumonia due to a known cause. This fraction was calculated for each year in 1999 – 2009, and averaged.

S. pneumoniae was taken to cause 15.4% of CAP, and that fraction was taken to be the same for nosocomial pneumonia. The fraction of CAP caused by *S. pneumoniae* was calculated to be almost identical between inpatients and outpatients in a large retrospective study [94]; therefore we took the fraction of 15.4% to apply to both inpatient and outpatient subjects.

The incidence of pneumococcal pneumonia in 1999 was then calculated to be 22,800, 38,000, and 489,200 cases, from youngest to oldest age group, respectively. This equated to

301.0, 69.4, and 232.5 cases per 100,000 people, based on the 1999 US population as shown in Table B.1.

Model-based estimates of fractions of subjects colonized with *S. pneumoniae* in 1999 (Table C.1) were used to determine the fraction of colonized subjects who developed pneumonia within each age group, prior to the introduction of the vaccine. A small fraction of subjects was estimated by the model to be dually colonized with both vaccine-type (VT) and non-vaccine-type (NVT) serotypes; these subjects were assumed to have twice the probability of developing pneumonia. The model structure did not allow subjects to be dually colonized with two VT serotypes, or dually colonized with two NVT serotypes.

The serotypes included in the vaccine were those that were most likely to cause IPD in the United States, making them likely to be more virulent than the non-vaccine serotypes. Therefore, case:colonization ratios were allowed to vary between VT and NVT serotypes in the dynamic transmission model. The fitted VT case:colonization ratios for reported IPD were 7.7, 1.8, and 1.03 times higher than the equivalent NVT case:colonization ratios, from youngest to oldest age group respectively (Chapter 3).

These relationships between VT and NVT case:colonization ratios were used to calculate case:colonization ratios for pneumococcal pneumonia, by serotype and age groups. This resulted in case:colonization ratios of 0.0017, 0.000088, and 0.00035 for VT pneumonia, from youngest to oldest age group respectively, and corresponding ratios of 0.000022, 0.000049, and 0.00034 for NVT pneumonia. These ratios were then used to estimate the number of cases of pneumococcal pneumonia each year.

As described in Chapter 3, fitting the model to the observed data resulted in an estimated 55% reduction in the chance of VT colonization due to vaccination. Vaccination has been shown to cause an overall 95% decrease in the acquisition of VT IPD [68; 69], so the vaccine was calculated to cause an 89% decrease in the development of VT IPD among vaccinated subjects who were colonized with VT serotypes. Therefore, to calculate the cases of VT

pneumonia among vaccinated children, the number of VT pneumonia cases among vaccinated people was estimated as 11% of the product of the population size, the case:colonization ratio, and the fraction of children who were vaccinated and had VT colonization, by year and age group. The numbers of cases of NVT pneumonia and VT pneumonia among non-vaccinated people were similarly calculated as the products of the population size, the case:colonization ratio, and the appropriate fractions of subjects. The final model-based estimates of pneumococcal pneumonia are presented in Table 5.3 by age group and year.

Table 5.3
Model-Estimated Pneumonia Incidence by Age Group

	Number of Cases (Rate [a])			
Year	<2 years	2-15 years	≥16 years	Overall
1999	24,100 (318.1)	37,800 (69.0)	491,900 (233.8)	553,800 (203.0)
2000	24,200 (291.4)	38,800 (68.5)	507,700 (233.1)	570,700 (201.8)
2001	9,900 (117.0)	33,300 (58.7)	448,800 (203.6)	492,000 (172.2)
2002	5,900 (69.0)	24,900 (43.7)	369,300 (165.6)	400,100 (138.7)
2003	5,300 (61.5)	19,800 (34.7)	327,000 (145.2)	352,100 (121.0)
2004	5,200 (59.4)	17,400 (30.5)	308,800 (135.5)	331,400 (112.9)
2005	5,000 (56.5)	16,300 (28.6)	302,400 (131.2)	323,700 (109.3)
2006	5,100 (57.3)	16,000 (28.1)	301,800 (129.4)	322,900 (107.9)
2007	5,200 (57.4)	15,800 (27.6)	303,600 (128.6)	324,600 (107.4)
2008	5,200 (56.7)	15,800 (27.6)	306,300 (128.4)	327,300 (107.3)
2009	4,800 (56.3)	15,900 (27.6)	309,100 (128.3)	329,800 (107.4)

a Rates indicated number of cases per 100,000 subjects, within each age group.

Pneumonia cases in which the infection spreads into the bloodstream are typically referred to as invasive pneumonia, but are still diagnosed as pneumonia. The incidence of pneumococcal pneumonia was calculated using the number of patients diagnosed with pneumonia, which was taken to include those pneumonia patients in which the bacteria have spread to the blood (bacteremia). Therefore, the fraction of reported IPD due to pneumonia with bacteremia was taken to be included in these estimates of pneumonia, and was not counted separately to avoid duplication.

Comparisons to the Hypothetical No-Vaccine Scenario

We hypothesized that, if no pneumococcal conjugate vaccine had been introduced, the rate of pneumococcal disease incidence (i.e., the number of cases per 100,000 subjects) would remain constant at 1999 levels, within age group and disease type. In this scenario, although the rates remained constant, the hypothetical number of pneumococcal disease cases per year increased due to corresponding increases in population size.

The numbers of pneumococcal disease cases estimated under this scenario are presented in Table 5.4. The resulting differences between the hypothetical numbers and the model-based estimates displayed in Tables 5.2 and 5.3, for each year and summed over the total ten-year period, are also presented in Table 5.4. The difference between these two estimates represented the estimated number of cases of pneumococcal disease prevented. We estimated that the vaccine has prevented 16,490 cases of meningitis, 70,500 cases of bacteremia, and more than 2.2 million cases of pneumonia in the United States since its introduction.

Table 5.4
Differences Between Estimated and Hypothetical Pneumococcal Disease Cases, Across All Age Groups, by Year

Year	Hypothetical Model-Estimated Cases in the Absence of Vaccine				Difference between Model-Estimated Numbers of Cases and Hypothetical Cases			
	Menin-gitis	Bacte-remia	Pneumonia	Overall	Menin-gitis	Bacte-remia	Pneumonia	Overall
1999	3,700	16,700	553,800	574,200	0	0	0	0
2000	3,870	17,500	574,500	595,870	170	700	3,800	4,670
2001	3,930	17,700	581,300	602,930	1,000	4,200	89,300	94,500
2002	3,970	17,900	587,500	609,370	1,540	6,600	187,400	195,540
2003	4,000	18,000	593,200	615,200	1,780	7,500	241,100	250,380
2004	4,040	18,300	599,600	621,940	1,900	8,300	268,200	278,400
2005	4,090	18,400	605,900	628,390	1,970	8,400	282,200	292,570
2006	4,130	18,600	612,800	635,530	2,010	8,600	289,900	300,510
2007	4,170	18,800	619,800	642,770	2,040	8,700	295,200	305,940
2008	4,210	19,000	626,500	649,710	2,060	8,800	299,200	310,060
2009	4,190	19,000	630,100	653,290	2,020	8,700	300,300	311,020
Total	...	...	...	...	16,490	70,500	2,256,600	2,343,590

Costs of Pneumococcal Disease

The costs of hospitalization for patients with pneumococcal isolates from the CSF, blood, or lungs were collected from the Duke hospital system during 1996 – 2006, and converted to 2010 dollars, as described in Chapter 4. The average cost of hospitalization for each type of infection was then used in conjunction with estimated disease incidence for each disease to calculate estimated disease costs. Patients with pneumococcal isolates were categorized hierarchically as CSF, blood, or lung, where subjects with both blood and lung isolates were categorized as blood, and subjects with isolates from both CSF and another location were categorized as CSF. Due to the limited number of patients with financial information, and the lack of a statistically significant relationship between age and cost, average costs for patients across all age groups were used.

The mean cost of hospitalization for subjects with pneumococcal isolates from the CSF was $15,864 (SE, $3667). Due to the severity of this disease, all subjects with pneumococcal meningitis were considered to have been hospitalized. The average mean hospitalization cost of subjects with pneumococcal isolates from the blood was somewhat higher than that of meningitis, at $24,417 (SE, $3447). Based on an earlier study of pneumococcal bacteremia [132], 97% of subjects with bacteremia were taken to be hospitalized, and the remaining subjects to be treated as outpatients. The average cost of hospitalization for patients with pneumococcal lung isolates was $48,558 (SE, $4154). Sixty-two percent of subjects with pneumococcal pneumonia were taken to be hospitalized, and the remaining 38% treated as outpatients, based on a large retrospective study of CAP [94].

The cost of outpatient treatment for pneumococcal disease has been very poorly characterized in the literature. Recent research [133] indicated that inpatient costs for all-cause CAP were 20-25 times the cost of outpatient treatment. Based on this, outpatient direct medical costs were taken to be 5% of hospitalization costs. Therefore, outpatient costs were calculated to be $1221 for bacteremia and $2428 for pneumonia.

Hospitalization costs typically do not include fees from physicians and other professionals, and hence underestimate the true cost of treatment. Physician fees for pneumonia were calculated to be 6.581% of hospitalization costs, based on a national sample of patients with CAP [119]. We took this ratio to be consistent for all types of pneumococcal disease, and increased all costs by this fraction to incorporate physician fees.

We combined these cost figures with model-based estimates of pneumococcal disease incidence (Tables 5.2 and 5.3) to calculate the annual treatment costs, in millions of dollars, as displayed in Table 5.5. Overall, the total estimated direct medical cost for pneumococcal disease for 1999 through 2009 was approximately 147 billion dollars.

Table 5.5
Estimated Annual Costs of Pneumococcal Disease, in Millions of Dollars

	Menin-	Bacteremia			Pneumonia			
Year	gitis	Inpatient	Outpatient	Total	Inpatient	Outpatient	Total	Overall
1999	$62.6	$421.6	$0.7	$422.2	$17,769.9	$544.6	$18,314.5	$18,799.3
2000	$62.6	$424.1	$0.7	$424.8	$18,312.2	$561.2	$18,873.4	$19,360.7
2001	$49.5	$340.8	$0.5	$341.3	$15,786.9	$483.8	$16,270.7	$16,661.6
2002	$41.1	$285.3	$0.4	$285.7	$12,838.1	$393.4	$13,231.5	$13,558.3
2003	$37.5	$265.1	$0.4	$265.5	$11,297.9	$346.2	$11,644.1	$11,947.1
2004	$36.2	$252.4	$0.4	$252.8	$10,633.7	$325.9	$10,959.6	$11,248.6
2005	$35.8	$252.4	$0.4	$252.8	$10,386.6	$318.3	$10,704.9	$10,993.6
2006	$35.8	$252.4	$0.4	$252.8	$10,361.0	$317.5	$10,678.5	$10,967.2
2007	$36.0	$255.0	$0.4	$255.4	$10,415.5	$319.2	$10,734.7	$11,026.1
2008	$36.4	$257.5	$0.4	$257.9	$10,502.2	$321.8	$10,824.0	$11,118.2
2009	$36.7	$260.0	$0.4	$260.4	$10,582.4	$324.3	$10,906.7	$11,203.8
Total	$470.2	$3,266.5	$5.1	$3,271.5	$138,886.4	$4,256.2	$143,142.6	$146,884.3

Note: Costs are displayed in millions of 2010 dollars.

The hypothetical pneumococcal disease costs in the absence of vaccination, based on the number of hypothetical cases displayed in Table 5.4, were similarly calculated. The prevented medical costs due to vaccination during 2000 – 2009 were estimated as the difference between costs based on model-estimated pneumococcal incidence and hypothetical costs assuming constant disease rates at pre-vaccination levels, as displayed in Table 5.6.

For the first ten years following the introduction of the vaccine, we estimated that 13.0 billion dollars were spent on the vaccine. That cost was subtracted from the prevented disease cost of 76.7 billion dollars, resulting in 63.7 billion dollars of cost savings. Over 95% of estimated cost savings were due to prevented pneumonia cases (which were the most poorly characterized), with approximately seven percent of estimated cost savings due to prevented bacteremia cases. Estimated cost savings due to prevented meningitis were less than one percent. We then predicted the future cost savings due to vaccination through year 2025.

Table 5.6
Hypothetical Pneumococcal Disease Costs In the Absence of Vaccine, and Differences, in Millions of Dollars

	Hypothetical Scenario, Without Vaccine				Difference between Hypothetical Scenario and Model-Based Estimates			
Year	Menin-gitis	Bacte-remia	Pneumonia	Total	Menin-gitis	Bacte-remia	Pneumonia	Total
1999	$62.6	$422.2	$18,314.5	$18,799.3	--	--	--	--
2000	$65.4	$442.4	$18,999.0	$19,506.9	$2.9	$17.7	$125.6	$146.2
2001	$66.4	$447.5	$19,223.9	$19,737.8	$16.9	$106.2	$2,953.2	$3,076.3
2002	$67.1	$452.6	$19,428.9	$19,948.6	$26.0	$166.9	$6,197.4	$6,390.3
2003	$67.6	$455.1	$19,617.4	$20,140.1	$30.1	$189.6	$7,973.3	$8,193.0
2004	$68.3	$462.7	$19,829.1	$20,360.1	$32.1	$209.8	$8,869.5	$9,111.5
2005	$69.2	$465.2	$20,037.5	$20,571.8	$33.3	$212.4	$9,332.6	$9,578.3
2006	$69.8	$470.3	$20,265.7	$20,805.8	$34.0	$217.4	$9,587.2	$9,838.6
2007	$70.5	$475.3	$20,497.2	$21,043.0	$34.5	$220.0	$9,762.5	$10,016.9
2008	$71.2	$480.4	$20,718.7	$21,270.2	$34.8	$222.5	$9,894.7	$10,152.0
2009	$70.8	$480.4	$20,837.8	$21,389.0	$34.2	$220.0	$9,931.1	$10,185.2
Total	$749.0	$5,053.9	$217,769.7	$223,572.6	$278.8	$1,782.4	$74,627.1	$76,688.3

Note: Costs are displayed in millions of 2010 dollars.

Predicted Impact of Vaccination Effects Through 2025

The estimated impact of vaccination was further evaluated over the next 16-year period (2010 – 2025). Population levels during that period were taken to be linear and predicted using linear regression, although the population growth over a longer period is more complicated. All linear regressions were performed on data from 1999 – 2009 (Appendix B), to obtain predictions for 2010 – 2025.

The predicted population size was first calculated within each age group, for years 2010 – 2025, using the data from 1999 - 2009. The ratio of births to number of adults (at least 16 years of age) was calculated for each year during 1999 – 2009. A linear regression was then applied to this birth rate to predict the birth rate in subsequent years. The predicted birth rate was multiplied by the predicted population of adults to predict the number of births each year. The infant death rate for 1999 – 2009 was similarly used to predict future infant death rates using a linear regression, and to predict the number of infant deaths for 2010 – 2025. The number of children eligible for vaccination was taken to be the predicted number of children surviving to 1 year of age. The percentages of eligible children receiving 1, 2, 3, or 4 doses were held constant at levels calculated for the 2006 birth cohort, as displayed in Table 5.1. The predicted number of doses administered each year was then calculated (Table B.2).

A new pediatric pneumococcal vaccine, PCV13, was approved in February 2010. This vaccine includes the seven serotypes present in the original PCV7 vaccine, plus six additional serotypes. Very limited information is publicly available regarding historic incidence of colonization and IPD caused by these six additional serotypes. Therefore, the existing model and its projections regarding the incidence of colonization and reported IPD were used for these future predictions. Consequently, these predictions did not incorporate the expected decreases in colonization and IPD due to the six serotypes not included in the PCV7 vaccine, leading to an underestimation of reductions in disease burden after the PCV13 introduction.

The estimates of vaccination cost incorporated the cost of the PCV13 vaccine, $108.75 in 2010 and $114.15 in 2011, and assumed that the vaccine cost remains constant in future years. Twelve dollars was also included for estimated vaccine administration costs. This increase in vaccination cost without any corresponding increase in the characterization of vaccine effects was again a conservative approach, leading to a possible overestimation of vaccine costs in relation to disease costs, and an underestimation of cost savings due to

vaccination. Over 28 billion dollars were predicted to be spent on pediatric pneumococcal vaccination in 2010 through 2025 (Table B.2).

The incidence of pneumococcal disease was predicted for the same period, using model-based estimates from the dynamic transmission model (Tables C.1 and C.2), as described in Chapter 3. The previously described methodology was used to determine the rates of meningitis, bacteremia, and pneumonia from those estimates. The fraction of estimated IPD attributable to meningitis was 5.35%, 6.22%, and 4.38%, from youngest to oldest age group. Bacteremia incidence was calculated as 21.4% of estimated IPD. Pneumococcal pneumonia incidence was calculated using model-based estimates of colonization, and case:colonization ratios by age group and serotype group. Predicted pneumococcal disease incidence and costs, in 2010 dollars, are presented in Table 5.7.

Table 5.7
Predicted Pneumococcal Disease Incidence and Associated Costs

	Number of Cases (Rate [a])			Predicted Costs (Millions of Dollars)			
Year	Meningitis	Bacteremia	Pneumonia	Menin-gitis	Bacte-remia	Pneumonia	Total
2010	2,200 (0.7)	10,400 (3.3)	334,600 (107.4)	$37	$263	$11,065	$11,366
2011	2,220 (0.7)	10,500 (3.3)	338,300 (107.6)	$38	$265	$11,188	$11,491
2012	2,240 (0.7)	10,600 (3.3)	341,900 (107.6)	$38	$268	$11,307	$11,613
2013	2,280 (0.7)	10,700 (3.3)	345,600 (107.8)	$39	$271	$11,429	$11,738
2014	2,300 (0.7)	10,800 (3.3)	349,300 (107.9)	$39	$273	$11,552	$11,864
2015	2,320 (0.7)	10,900 (3.3)	353,000 (108.0)	$39	$276	$11,674	$11,989
2016	2,340 (0.7)	11,100 (3.4)	356,700 (108.1)	$40	$281	$11,796	$12,116
2017	2,370 (0.7)	11,200 (3.4)	360,300 (108.2)	$40	$283	$11,915	$12,239
2018	2,390 (0.7)	11,300 (3.4)	364,100 (108.4)	$40	$286	$12,041	$12,367
2019	2,410 (0.7)	11,400 (3.4)	367,700 (108.4)	$41	$288	$12,160	$12,489
2020	2,440 (0.7)	11,500 (3.4)	371,500 (108.6)	$41	$291	$12,286	$12,618
2021	2,460 (0.7)	11,600 (3.4)	375,200 (108.7)	$42	$293	$12,408	$12,743
2022	2,480 (0.7)	11,700 (3.4)	378,900 (108.8)	$42	$296	$12,531	$12,868
2023	2,500 (0.7)	11,800 (3.4)	382,600 (108.9)	$42	$298	$12,653	$12,993
2024	2,530 (0.7)	11,900 (3.4)	386,300 (109.0)	$43	$301	$12,775	$13,119
2025	2,550 (0.7)	12,000 (3.4)	390,000 (109.1)	$43	$303	$12,898	$13,244
Total	...	...	...	$643	$4,536	$191,677	$196,856

Note: Costs are displayed in millions of 2010 dollars.
a Rates indicated number of cases per 100,000 subjects, within each age group.

The hypothetical scenario of the vaccine not being introduced was then expanded to this time period, by holding the rate of pneumococcal disease, within age group and disease type, constant at pre-vaccination levels. The predicted disease burden in this hypothetical scenario, and the difference in direct medical costs between these hypothetical costs and model-based predictions of costs, are displayed in Table 5.8.

More than 196.8 billion dollars in direct medical costs were predicted to have been prevented during these 16 years, with a cost savings of 168.5 billion dollars after accounting for vaccination costs. Overall, the pneumococcal vaccine was predicted to prevent at least 10.3 billion dollars in pneumococcal disease costs each year during this period.

Table 5.8
Predicted Hypothetical Pneumococcal Disease Costs In the Absence of Vaccine, and Differences, in Millions of Dollars

	Hypothetical, Without Vaccine				Difference between Hypothetical and Actual			
Year	Menin-gitis	Bacte-remia	Pneu-monia	Total	Menin-gitis	Bacte-remia	Pneu-monia	Total
2010	$73	$493	$21,182	$21,748	$36	$230	$10,116	$10,382
2011	$74	$498	$21,413	$21,985	$36	$233	$10,225	$10,494
2012	$74	$503	$21,641	$22,219	$36	$235	$10,335	$10,606
2013	$75	$508	$21,876	$22,459	$37	$238	$10,447	$10,721
2014	$76	$513	$22,104	$22,693	$37	$240	$10,553	$10,830
2015	$77	$521	$22,339	$22,937	$37	$245	$10,665	$10,948
2016	$77	$523	$22,571	$23,171	$38	$243	$10,774	$11,055
2017	$78	$528	$22,799	$23,405	$38	$245	$10,884	$11,167
2018	$79	$536	$23,030	$23,645	$39	$250	$10,989	$11,278
2019	$80	$539	$23,262	$23,880	$39	$250	$11,102	$11,391
2020	$80	$546	$23,487	$24,113	$39	$255	$11,201	$11,495
2021	$81	$549	$23,725	$24,355	$40	$255	$11,317	$11,612
2022	$82	$554	$23,953	$24,589	$40	$258	$11,423	$11,720
2023	$83	$561	$24,185	$24,828	$40	$263	$11,532	$11,835
2024	$83	$564	$24,416	$25,063	$41	$263	$11,641	$11,944
2025	$84	$571	$24,644	$25,300	$41	$268	$11,747	$12,056
Total	$1,255	$8,507	$366,627	$376,390	$612	$3,972	$174,950	$179,534

Note: Costs are displayed in millions of 2010 dollars.

Supplemental Analysis Using Costs Based on a Nationwide Sample

The hospitalization costs used in the previous sections were based on a single hospital system, and may not be representative of costs in other areas and other hospital systems.

Those costs were calculated as the average direct medical costs of patients admitted to the hospital who had a *S. pneumoniae* isolate collected from the CSF, blood, or lower respiratory tract, as described in Chapter 4. Therefore, those costs included patients whose primary diagnosis was not pneumococcal disease. A supplemental analysis was performed using costs from a national sample of subjects, based on primary diagnosis.

Costs from subjects with a primary diagnosis of pneumococcal meningitis (949 patients), all-cause bacteremia (26,939 patients), or pneumococcal pneumonia (26,736 patients) were obtained from the 2009 HCUP survey [96]. Costs for the three age groups used in the model were calculated from six age groups (<1, 1-17, 18-44, 45-64, 65-84, and 85+ years), and converted to 2010 dollars. Costs were further increased by 6.581% to account for physician fees [119].

The average cost for subjects with pneumococcal meningitis (ICD-9-CM code 320.1) was $26,933 for young children, $20,401 for older children, and $23,722 for adults (SE, $3611, $3004, $1453 respectively). The average cost for subjects with all-cause bacteremia (ICD-9-CM code 790.7) was $9342, $13,311, and $12,838 (SE, $1959, $3544, $635) from youngest to oldest age group, respectively. Subjects with pneumococcal pneumonia (ICD-9-CM code 481) had average costs of $13,006 (SE, $1688) for children less than 16 years of age, and $11,439 (SE, $461) for patients at least 16 years of age. These costs were combined with model-based estimates of disease incidence (Tables 5.2 and 5.3) to estimate medical costs of pneumococcal disease during 1999 – 2009, as displayed in Table 5.11.

The hypothetical scenario of vaccine not being introduced was again applied in this supplemental analysis, by holding the disease rates constant at pre-vaccination levels, within age group and disease type. Model-based estimates of pneumococcal disease cases prevented (Table 5.4) were used to calculate estimated cost savings, by disease type and overall (Table 5.9).

Table 5.9
Estimated Annual Costs of Pneumococcal Disease using HCUP Costs, and Differences, in Millions of Dollars

	Estimated Annual Costs				Difference between Hypothetical Scenario and Model-Based Estimates			
Year	Menin-gitis	Bacteremia	Pneumonia	Total	Menin-gitis	Bacteremia	Pneumonia	Total
1999	$89.0	$839.7	$4,110.0	$5,038.7	$0.0	$0.0	$0.0	$0.0
2000	$88.8	$844.7	$4,233.9	$5,167.4	$4.4	$35.2	$31.0	$70.6
2001	$69.2	$678.8	$3,637.9	$4,385.8	$25.5	$211.2	$677.3	$914.0
2002	$57.0	$568.2	$2,953.8	$3,578.9	$38.7	$331.8	$1,407.1	$1,777.7
2003	$52.0	$527.9	$2,598.1	$3,178.1	$44.4	$377.1	$1,804.8	$2,226.3
2004	$50.2	$502.8	$2,445.0	$2,998.0	$47.3	$417.3	$2,005.1	$2,469.7
2005	$49.7	$502.8	$2,387.3	$2,939.7	$49.0	$422.3	$2,109.1	$2,580.4
2006	$49.7	$502.8	$2,381.5	$2,934.0	$49.9	$432.4	$2,165.6	$2,647.9
2007	$49.9	$507.8	$2,393.7	$2,951.4	$50.7	$437.4	$2,205.1	$2,693.2
2008	$50.4	$512.8	$2,413.5	$2,976.7	$51.2	$442.5	$2,234.8	$2,728.4
2009	$50.9	$517.9	$2,431.5	$3,000.2	$50.0	$437.4	$2,241.1	$2,728.5
Total	$656.7	$6,506.1	$31,986.2	$39,149.0	$411.2	$3,544.6	$16,881.0	$20,836.8

Note: Costs are displayed in millions of 2010 dollars.

This supplemental analysis estimated that 39.1 billion dollars were spent on pneumococcal disease-related medical costs during 1999 - 2009, which is significantly less than the 146.8 billion dollars estimated by the original analysis. This analysis further estimated that 20.8 billion dollars in medical costs were prevented during the first ten years of vaccination, as compared to the 76.7 billion dollars in prevented costs previously estimated by the original analysis.

This large difference between estimates of prevented costs was driven by differences in estimates of pneumonia costs. The estimated cost per subject ranged from $51,574 for data from the Duke hospital system (including estimated physician costs) to $11,439 - $13,006, depending on age, for HCUP estimates. In spite of the large differences in estimated pneumonia costs, the PCV7 vaccine is still considered cost-effective using estimates from the HCUP data, with estimated cost savings of 7.8 billion dollars over the first ten years.

The predicted impact of vaccination during 2010 – 2025 was also evaluated using HCUP costs. Model-based estimates of pneumococcal disease (Table 5.7) were again compared to the hypothetical no-vaccine scenario. Predicted medical costs, and predicted cost savings

due to vaccination, are displayed in Table 5.10. These estimates reflect a predicted 52.6 billion dollars in pneumococcal disease costs, and 48.3 billion dollars in prevented medical costs, during this 16-year period. After incorporating the 28.3 billion dollars in predicted vaccination costs for that period, 19.9 billion dollars were predicted to be saved during 2010 – 2025, in addition to the 7.8 billion dollars in cost savings estimated during the first ten years of vaccination.

Table 5.10
Predicted Hypothetical Pneumococcal Disease Costs In the Absence of Vaccine, and Differences, using HCUP Costs, in Millions of Dollars

	Predicted Costs				Difference between Hypothetical and Actual			
Year	Menin-gitis	Bacte-remia	Pneumonia	Total	Menin-gitis	Bacte-remia	Pneumonia	Total
2010	$52	$523	$2,467	$3,041	$52	$458	$2,284	$2,794
2011	$52	$528	$2,494	$3,074	$53	$463	$2,309	$2,824
2012	$53	$533	$2,521	$3,106	$53	$468	$2,333	$2,854
2013	$53	$538	$2,548	$3,139	$54	$473	$2,358	$2,885
2014	$54	$543	$2,575	$3,172	$54	$478	$2,382	$2,914
2015	$54	$548	$2,602	$3,205	$55	$488	$2,407	$2,950
2016	$55	$558	$2,629	$3,242	$55	$483	$2,432	$2,970
2017	$56	$563	$2,656	$3,274	$56	$488	$2,456	$2,999
2018	$56	$568	$2,684	$3,308	$57	$498	$2,480	$3,034
2019	$57	$573	$2,710	$3,339	$57	$498	$2,505	$3,060
2020	$57	$578	$2,738	$3,373	$57	$508	$2,527	$3,092
2021	$58	$583	$2,765	$3,406	$58	$508	$2,553	$3,119
2022	$58	$588	$2,792	$3,439	$59	$513	$2,577	$3,148
2023	$59	$593	$2,819	$3,471	$59	$523	$2,601	$3,183
2024	$59	$598	$2,846	$3,504	$60	$523	$2,625	$3,208
2025	$60	$603	$2,874	$3,537	$60	$533	$2,649	$3,243
Total	$892	$9,020	$42,719	$52,631	$900	$7,899	$39,477	$48,276

Note: Costs are displayed in millions of 2010 dollars.

Discussion

We conclude that the pediatric pneumococcal vaccine introduced in 2000 has been significantly cost-effective in the first ten years of use, saving at least 7.8 billion dollars. We predict that it will continue to have significant cost savings in subsequent years, if the vaccine use remained constant. However, the benefits of the vaccine were not limited to the monetary aspect.

Mortality and Morbidity

In addition to summarizing estimated monetary costs, we also explored the model-based estimates of fatalities associated with pneumococcal disease. Mortality rates for pneumococcal meningitis, based on ABCs data as reported by Hsu [27], were estimated to be 8.4% among young children and older children, and 22.3% among adults 16 years or older. Mortality rates among patients with pneumococcal bacteremia were taken to be 3.5% among young children, 1.7% among older children, and 12.6% among adults [134]. Due to the lack of specific information on outpatient mortality and the small fraction (3%) of subjects not hospitalized [132], death rates were assumed to be the same among hospitalized and non-hospitalized subjects with pneumococcal bacteremia. Mortality rates among patients with pneumococcal pneumonia were taken to be 7.7% among hospitalized patients and 0.5% among outpatients [94].

These fractions were used in conjunction with estimates of prevented pneumococcal disease from our baseline model to estimate the number of deaths caused by the disease, and similarly estimate the number of deaths prevented by the vaccine. In 1999, an estimated 29,800 deaths were caused by pneumococcal disease, with 700 due to meningitis, 1700 due to bacteremia, and 27,000 due to pneumonia. The estimated annual deaths caused by pneumococcal disease decreased to 18,000 in 2009, with an estimated 120,000 deaths prevented by the vaccine during 2000 - 2009. We further predict that 280,800 additional deaths will be prevented during 2010 – 2025.

Patients with meningitis are also known to experience long-term side effects. Among survivors, approximately 32% of children less than 16 years old [135] and 26% of adults [103] experience hearing loss, and 5% of children [135] and 22% of adults [103] experience neurological sequelae. Therefore, this model estimated that the vaccine also prevented almost 4000 cases of hearing loss and 1700 cases of neurological impairment during the ten-year period.

A cost-benefit analysis was chosen as the most appropriate type of cost-effectiveness analysis, as the main question of interest is whether the vaccine is of financial benefit, i.e. whether the estimated cost savings exceeded the estimated vaccine expenditures. Recent cost-effectiveness analyses [62; 64] have evaluated vaccination cost per life-year saved or cost per pneumococcal disease episode prevented. Other authors [63] have conducted a cost-utility analysis, whose primary outcomes were life-years and quality-adjusted life-years gained. While we agree that reduced mortality is a valuable result of the vaccination, approximately 5% of pneumococcal disease cases result in death. Pneumococcal disease also results in very little long-term disability, with 1266 estimated cases (779 hearing loss; 487 neurological impairment) in 1999. Due to the relatively low fatality and disability rates, we felt that monetary savings were of greater interest than mortality and morbidity.

Effects of Vaccination, by Age Group

The pediatric pneumococcal vaccine has had the biggest impact on pneumococcal disease in children less than two years of age. Our model-based estimates suggested that incidence of pneumococcal meningitis decreased from 10.3 to 0.2 cases per 100,000 children under 2 years of age, with concurrent decreases in pneumococcal bacteremia (40.8 to 0.9 cases per 100,000 children) and pneumococcal pneumonia (318.1 to 56.3 cases per 100,000 children) from 1999 to 2009. Over this ten-year period, the model estimated that 7700 cases of meningitis, 30,600 cases of bacteremia, and 201,500 cases of pneumonia were prevented in children less than 2 years of age. This equated to an approximate 60% reduction in each type of pneumococcal disease within this age group, and a savings of 7.6 billion dollars in pneumococcal disease costs among young children.

The model characterized the effect of vaccination as persisting through 15 years of age, and waning as children move into the adult age group (≥ 16 years old). The disease burden per 100,000 subjects in older children (2 – 15 years old) was the lowest of all three age groups prior to the introduction of the vaccine. Therefore, smaller decreases were seen in this age

group than in young children. Model-based estimates indicated that pneumococcal meningitis incidence among older children decreased from 0.7 to 0.4 cases per 100,000 children, with pneumococcal bacteremia incidence decreasing from 2.3 to 1.2 cases per 100,000 children, and pneumococcal pneumonia decreasing from 69.0 to 27.6 cases per 100,000 children. The model estimated that these decreases resulted in the prevention of 1480 cases of meningitis, 4800 cases of bacteremia, and 178,800 cases of pneumonia in children 2 – 15 years of age during the first ten years of vaccination. This was associated with a 6.1 billion dollar cost savings among older children.

The model-estimated reductions in pneumococcal disease among adults 16 years of age and older was a reflection of the decreases in colonization, and additional decreases in pneumococcal disease, due to herd immunity. Although substantial pneumococcal disease costs remained in this population, the model estimated that 63.1 billion dollars in direct medical costs among adults were prevented in the first ten years of pediatric vaccination. The model-estimated decreases in pneumococcal disease incidence among adults (1.2 to 0.8 meningitis cases per 100,000 adults; 5.9 to 3.9 bacteremia cases per 100,000 adults; and 233.8 to 128.3 pneumonia cases per 100,000 adults) during this ten year period, and the associated cost savings, demonstrated that the indirect effects of vaccination play an important part in the calculation of the cost-benefit of the vaccine. We acknowledge that, by conservatively not characterizing the pediatric vaccine as having any effect on subjects 16 years of age or older, the effects of this vaccine on adults was potentially underestimated.

Recent Literature

Huang et al. [136] evaluated the estimated pneumococcal disease burden in 2004, and included non-invasive pneumococcal infections (such as acute otitis media and sinusitis) and indirect costs (work-loss, lost wages, long-term disability costs) in the evaluation. The authors used various national surveys, such as the National Hospital Discharge Survey and the National Ambulatory Medical Care Survey, to estimate the incidence of these infections.

These calculations were based on primary diagnoses, which can underestimate the true incidence. Thirty percent of inpatient pneumonia, 20% of outpatient pneumonia, and 10-60% of meningitis (depending on age) were assumed to be caused by *S. pneumoniae*.

While certain estimates differ between our studies, we agree with the authors' findings that a significant disease burden existed in 2004 (and in subsequent years), with the majority of healthcare resource utilization due to pneumococcal disease occurring in adults. Of the 11.2 billion dollars in pneumococcal disease costs estimated by our model for 2009, 97% were due to pneumonia. We further calculated that 83% of those estimated pneumonia costs were due to adults. Further decreases in pneumococcal disease among adults, especially pneumonia, would greatly decrease the pneumococcal medical costs in the overall population.

Model Assumptions

We acknowledge that this cost-benefit analysis made certain assumptions. Our results were based on estimates of reported IPD and colonization from the dynamic transmission model described in Chapter 3. Any inaccuracies in those model-based estimates would be reflected in this analysis. Estimates of direct medical costs, as recorded by the Duke hospital system and described in Chapter 4, were assumed to be representative of average medical costs for subjects hospitalized with pneumococcal disease. These costs may have varied geographically and between different hospital systems. Due to poorly characterized outpatient direct medical costs, outpatient costs were taken to be 5% of inpatient costs [133], for both bacteremia and pneumonia.

The focus of this cost-benefit analysis was limited to the three most common types of pneumococcal disease: meningitis, bacteremia, and pneumonia. *S. pneumoniae* causes other illnesses, such as acute otitis media, sinusitis, and bone and joint infections [136]. Acute otitis media and sinusitis are typically treated on an outpatient basis, so were not included in the cost analyses presented in Chapter 4. Pneumococcal bone and joint infections are

infrequent, with only four such subjects identified during the ten-year data collection period at the Duke hospital system. Other types of infection, such as the isolates of *S. pneumoniae* from the eye or abdomen as identified in Chapter 4, are also very infrequent. Our analysis focused on the types of infection leading to the greatest medical costs (i.e., those which were most prevalent and which were primarily treated on an inpatient basis). We acknowledge that these cost estimates did not include all possible types of illness caused by *S. pneumoniae*. The decrease in colonization following the introduction of the vaccine was expected to cause decreases in these other types of infections, making the numbers presented here an underestimate of the true cost savings due to the vaccine.

The incidence of pneumococcal pneumonia was poorly characterized, and was calculated as a fraction of model-estimated colonized subjects, as described earlier. We acknowledge that the resulting estimates of pneumococcal pneumonia incidence were very sensitive to slight changes to parameter values used in these calculations. Where possible, we conservatively underestimated the disease burden. For example, we assumed that 15.4% of pneumonia was caused by *S. pneumoniae*, based on the 1999 – 2009 average fraction of pneumonia due to a known cause that is due to *S. pneumoniae*, and used that fraction in conjunction with 1999 colonization estimates to estimate case:colonization ratios. That average was lower than the 1999 fraction of 18.2% of pneumonia due to a known cause being due to *S. pneumoniae*. A large retrospective study found *S. pneumoniae* to be the cause in 31% of CAP cases with known etiology [94]. A meta-analysis of six large-scale studies of CAP found *S. pneumoniae* to be the most commonly identified cause of the disease [97]. Another study explored the use of lung aspirates in addition to standard microbiological procedures, and found that *S. pneumoniae* was present in 33% of cases whose etiology was not identifiable by normal means, and 30% of total cases [98]. Therefore, the incidence of pneumococcal pneumonia may easily be twice as high as was estimated in this analysis.

In addition, the estimated cost savings were so large that the pneumococcal vaccine would still be cost-effective if reductions in pneumonia costs were as low as 15% of the estimated reductions in the base case scenario. In other words, the vaccine would still be considered cost effective across very different characterizations of pneumonia effects. Therefore, although these calculations of pneumococcal pneumonia costs were based on a poorly characterized process, the conclusion that the vaccine is cost-effective was very robust to differences in pneumonia characterization.

Supplemental analyses were also performed using costs from a nationwide inpatient sample. These figures had the advantage of being nationally representative, instead of being limited to a single hospital system or geographic area, as well as being based on a much larger sample of patients. These costs were limited to subjects with a specific primary diagnosis, so did not include all subjects with a certain disease (for example, an immunocompromised subject who was admitted to the hospital with pneumonia might have the immunocompromising condition listed as the primary diagnosis, and pneumonia as a secondary diagnosis). Subjects who were assigned an unspecific diagnosis (such as pneumonia due to an unspecified organism) or incorrect diagnosis (such as streptococcal septicemia) were also excluded [138]. Furthermore, information was not available to determine what fraction of subjects with bacteremia developed septicemia (a more severe form of blood infection); average hospitalization costs for pneumococcal septicemia were close to twice the costs of bacteremia [96]. Therefore, these figures may have underestimated actual costs. In spite of differences in hospitalization costs between the Duke hospital system and the nationwide inpatient sample, the estimated cost savings during the first ten years was still significant, confirming the conclusion of cost-effectiveness of the vaccine.

Conclusion

Our cost-benefit analysis, incorporating a dynamic transmission model to estimate pneumococcal disease incidence by disease type, age group, and year, found that the PCV7 pediatric pneumococcal vaccine was significantly cost-effective within the first ten years. Our baseline model estimated a 63.7 billion dollar cost savings during this period, or approximately $499 saved for every vaccine dose administered. The model also predicted a 151.2 billion dollar cost savings during the next 16 years. The supplemental analysis, incorporating cost estimates from a nationwide inpatient sample, resulted in reduced estimates of cost savings, but also concluded that the vaccine was cost-effective. Due to the introduction of an expanded PCV13 pneumococcal vaccine in 2010, and the anticipated decreases in colonization and pneumococcal disease due to the six serotypes included in that vaccine but not in the PCV7 vaccine, we underestimated the cost savings for this later period.

6. CONCLUSIONS

Streptococcus pneumoniae is a leading cause of pneumonia, bacteremia, and meningitis. It is primarily transmitted by healthy, asymptomatically colonized subjects. A small fraction of colonized subjects develops pneumococcal disease. The Centers for Disease Control and Prevention's Active Bacterial Core surveillance (CDC ABCs) team tracks the incidence of invasive pneumococcal disease (IPD), including meningitis, bacteremia, and invasive pneumonia. Non-invasive pneumococcal pneumonia, in which the bacteria have not spread to the bloodstream, is also very prevalent but poorly characterized in the literature.

The PCV7 pneumococcal conjugate vaccine was approved for use in the United States in February 2000. Since that time, decreases in IPD have been observed across all age groups, with the largest decreases among young children. We developed a dynamic transmission model, which reflects the changes in colonization and IPD in the United States since the introduction of the vaccine, and predicts future colonization prevalence and IPD incidence. We also explored the relationship between antimicrobial resistance and inpatient length of stay (LOS), cost, and mortality in subjects with pneumococcal infections. We then combined the average hospitalization cost by isolate location with results from the dynamic transmission model, to evaluate the cost-effectiveness of the vaccine.

Dynamic Transmission Model

We developed an age-structured dynamic transmission model for *S. pneumoniae*, calibrated to observed IPD incidence. Subjects were categorized as young children (< 2 years old), older children (2 – 15 years old), and adults (16 years or older). This model estimated both colonization prevalence and IPD incidence, reflecting the changes due to the introduction of the vaccine, and predicted future levels of colonization and IPD through 2025. It used parameter values from recent literature, where possible, and fitted the remaining parameter values using observed data.

Estimates of parameter values for transmission probabilities between subjects and the average duration of temporary immunity are not available in the literature. Values for these parameters were obtained by fitting five model outcomes (fractions of colonized subjects within age group, fraction of young children colonized by vaccine-type [VT] serotypes, and time to recolonization in young children) to observed data. These fitted parameter values were: transmission probabilities of $p_1 = 0.022$, $p_2 = 0.0013$, and $p_3 = 0.00031$ contact^{-1} for VT serotypes, from youngest to oldest age group, respectively; transmission probabilities of $p_1 = 0.035$, $p_2 = 0.0021$, and $p_3 = 0.00050$ contact^{-1} for non-vaccine-type (NVT) serotypes, based on a ratio of 1.6 between NVT and VT transmission probabilities; and a 3.4 week average duration of temporary immunity.

Our model estimated that VT colonization prevalence decreased to approximately two percent of pre-vaccination levels within each age group during the first ten years. VT colonization was predicted to further decrease during the subsequent 16 years, falling below 0.1% of subjects in every age group by 2017. In other words, the prevalence of VT colonization in young children was estimated to decline from 2.2 million children in 1999 to less than 20 estimated children in 2005, with prevalence in the overall population decreasing from 19.3 million people to less than 1000 in 2005 and less than 20 in 2009.

The fraction of subjects colonized with NVT serotypes was estimated to increase by nine percent within each age group within the first ten years, and to remain approximately constant during the next 16-year period, with increases of less than 0.1% of subjects within any age group. This equated to an increase from 19.9 million subjects colonized with NVT serotypes in 1999 to 24.7 million subjects colonized in 2010. Although the fraction of subjects colonized was predicted to remain approximately constant after that point, a predicted 27.5 million subjects will be colonized with NVT serotypes in 2025, due to projected increases in population size.

Vaccine-type IPD was also estimated to decrease dramatically during the first ten years of vaccination. The largest changes were estimated to occur among young children, with incidence decreasing from 188.5 cases per 100,000 subjects, or 14,280 subjects in 1999, to virtual eradication by 2010. Similar decreases were also estimated to occur among older children, with VT IPD incidence decreasing from 5.9 cases per 100,000 subjects, or 3210 cases in 1999 to virtual eradication by 2010. Due to the effects of herd immunity, decreases in VT IPD incidence were also estimated to occur among adults, from 13.2 cases per 100,000 subjects, or 40,040 in 1999, to less than ten cases estimated in 2010.

The incidence of NVT IPD was estimated to grow by approximately nine percent within each age group during the first ten years of vaccination. This equates to increases from 3.8 to 4.1 cases per 100,000 young children, from 5.2 to 5.7 cases per 100,000 older children, and from 16.9 to 18.5 cases per 100,000 adults. While these numbers may seem quite small, these increases, combined with population growth, equated to an increase of 10,200 cases of NVT IPD annually by the end of 2010.

NVT IPD incidence per 100,000 subjects was predicted to remain approximately constant from 2010 through 2025, while VT IPD was predicted to virtually disappear by 2010. The total number of NVT IPD cases each year was estimated to increase due to projected increases in population size. Overall, the annual incidence of total IPD was estimated to decrease from 78,800 cases in 1999 to 47,000 cases in 2006, with slightly increases in subsequent years, leading to 56,800 estimated cases of NVT IPD occurring in 2025.

The associated sensitivity analysis provided insight into which parameter values had the greatest impact on model-estimated colonization and IPD levels. These findings provide direction for future research that may be conducted to refine the most critical parameter values. Model estimates were most strongly affected by contact rates and average duration of NVT colonization among young children, transmission probabilities among children less than 16 years old, and the relationship between VT and NVT transmission probabilities.

Improved estimates of these parameter values could be obtained using the following methods. Contact rates between young children could be further explored by surveying parents of young children. The average duration of colonization could be estimated by conducting a longitudinal study, collecting nasopharyngeal swabs from children on a weekly basis for a minimum of several months. The transmission probabilities, while not able to be explicitly observed, could be calculated from the longitudinal study by concurrently collecting information on average contacts, and calculating transmission probabilities from the infectious contact rate.

Estimates from our model differed somewhat from those from similar models developed by Temime et al. [44] and Melegaro et al. [46]. The major reasons for these differences were the characterization of vaccine effects and competition between serotypes among the different models. The Temime et al. model was published before large-scale information was available on vaccine effects, and characterized vaccination as providing 100% protection against VT colonization. It further assumed that the basic reproductive numbers were the same for both serotype groups ($R_{01} = R_{02}$), and that dual colonization was not possible. That model estimated that VT colonization would approach zero after 10 years, and that NVT colonization would increase by approximately 300% during that time.

The Melegaro et al. model, calibrated to observed data in the United Kingdom, characterized vaccination as reducing the probability of being colonized by VT serotypes by 75.6%, and that colonization by VT serotypes led to a 50% reduction in the chance of colonization with NVT serotypes. Their model estimated that VT IPD would approach eradication within 8 – 12 years, depending on the vaccination strategy, and that the total number of NVT IPD cases each year would increase by approximately 20% during that time, and then remain steady. The Snedecor et al. [47] model did not distinguish between VT and NVT serotypes, and characterized forces of infection in an apparently ad-hoc method. It did characterize vaccination as decreasing the probability of VT colonization and subsequently decreasing the

probability of VT IPD. It also exhibited the same pattern of estimates of total IPD as shown in our model. However, it was only intended to apply to the first ten years following the introduction of the vaccine, and so should not be used to predict future colonization or IPD incidence.

None of these previously published models was intended to predict future IPD incidence in the United States. Both the Temime et al. and Melegaro et al. models exhibited a high degree of competition between serotypes, leading to increased serotype replacement, which may not be suitable for this population [71]. The Melegaro et al. model also predicted that the total annual incidence of NVT IPD would stabilize approximately ten years after the introduction of the vaccine, which did not incorporate the anticipated increases in the American population size.

We developed an age-structured model that incorporates a two-part characterization of vaccine effect, incorporating both a reduction in VT colonization and an additional reduction in the chance of developing VT IPD, as supported by multiple studies [31-39; 68; 69]. It was calibrated to eleven years of reported IPD incidence [17], recently published information on the fractions of reported IPD due to VT and NVT serotypes [73] and recently published estimates of competition parameters [70]. By incorporating the most up-to-date information available, it estimates both colonization prevalence and IPD incidence in the United States since the introduction of the vaccine, and predicts future colonization and IPD.

Antimicrobial Resistance and Inpatient Cost, Length of Stay, and Mortality

We evaluated the effect of antimicrobial resistance on hospitalization cost, LOS, and inpatient mortality in patients who had *S. pneumoniae* bacteria isolated from the lower respiratory tract or another normally sterile site. Patients were classified hierarchically into cerebrospinal fluid (CSF; $n = 32$), blood ($n = 265$), lung ($n = 267$), or 'other sterile site' ($n = 14$) categories, according to isolate location. Patients with isolates from multiple locations were classified as CSF, then blood, then lung, then other sterile site. Less than half

of the 578 patients exhibited resistance to one or more classes of antimicrobial agents, with the most frequent resistance found against sulfonamides (47.1% of isolates tested) and beta-lactams (42% of isolates tested). Fifty-three percent of patients were susceptible to all antimicrobial agents they were tested against, and 28% of patients exhibited multi-drug resistance (MDR; resistance to 3 or more classes of antimicrobial agents).

Eighteen percent of patients died during hospitalization; no significant relationship was found between mortality and either antimicrobial resistance or isolate location. LOS was significantly related to both isolate location and resistance, and cost was significantly related to isolate location, race, and resistance. Those relationships were present for both dichotomous resistance, in which patients were classified as susceptible or resistant to one or more classes of antimicrobial agents, and incremental resistance, in which patients were classified as resistant to 0, 1, 2, or 3 or more classes of antimicrobial agents.

LOS was consistently higher among subjects demonstrating resistance to at least one class of antimicrobial agents, after controlling for isolate location. Subjects exhibiting MDR had consistently longer estimated LOS than those with resistance to one or two classes of antimicrobial agents, after controlling for isolate location. Patients categorized as having lung isolates had the longest average LOS.

Cost analyses were performed on a subset of 274 patients. Isolate location, race, resistance, and race by resistance interaction were significantly related to hospitalization costs. Black patients who were susceptible to all antimicrobial agents tested had higher costs than black patients exhibiting resistance to at least one class of antimicrobials. That pattern was reversed for white patients and patients of other races, who had higher costs among subjects exhibiting resistance to at least one class of antimicrobials. Costs were highest among subjects with *S. pneumoniae* isolated from the lung, but not from the blood or CSF.

The longest estimated LOS and highest estimated costs were associated with patients with lung infections. Estimated LOS was consistently higher among resistant patients, as were

estimated costs among white patients and those of 'other' races. Estimated costs among black patients were higher among susceptible patients. Patients with multi-drug resistance consistently had the highest estimated costs.

Cost-Effectiveness of PCV7 Vaccine

The cost effectiveness of the PCV7 pneumococcal conjugate vaccine was evaluated using a cost-benefit analysis. The estimated annual incidences of pneumococcal meningitis, bacteremia, and pneumonia were calculated using model-based estimates of IPD and colonization. Direct medical costs for these three types of pneumococcal disease were estimated each year. Vaccination costs were calculated based on the estimated number of doses administered, and cost savings due to vaccination subsequently estimated.

The number of disease cases prevented, by age group and disease type, was calculated as the difference between these estimates of disease burden and the estimated disease burden under a hypothetical scenario in which the number of cases per 100,000 subjects was held constant at pre-vaccination levels. Average hospitalization costs were assigned to each type of disease, and the amount of prevented disease cost totaled over the first ten years of vaccination. The total cost savings during that time was calculated as the prevented disease cost minus the cost of vaccination during the first ten years. We concluded that the PCV7 vaccine has been cost-effective in the United States, with estimated cost savings of 7.8 to 63.7 billion dollars during the first ten years.

Cost-effectiveness was assessed in two ways. First, average hospitalization costs for patients with systemic pneumococcal infections who were admitted to the Duke hospital system (Chapter 4) were used. As a supplemental analysis, hospitalization costs from a nationwide inpatient sample, by age group and disease type, were also used [96].

An estimated 16,490 cases of pneumococcal meningitis, 70,500 cases of pneumococcal bacteremia, and 2,256,600 cases of pneumococcal pneumonia were prevented during the first ten years of vaccination, using the baseline model. After incorporating a 6.581% increase to

account for physician costs [119], the average hospitalization costs of subjects with pneumococcal isolates were $16,908 (CSF), $26,024 (blood), and $51,754 (lung), based on the Duke hospital system. These costs were based on all subjects with an isolate from that location, regardless of diagnosis. These average costs therefore included patients who had a primary diagnosis other than pneumococcal disease. In addition, these costs were from a single hospital system, and may not be representative of nationwide costs. Using these costs, an estimated 63.7 billion dollars was saved during the first ten years of vaccination, with an additional 151.2 billion dollars in predicted cost savings through 2025.

The supplemental analysis was conducted using costs from a nationwide inpatient sample, from the Healthcare Cost and Utilization Project (HCUP) [96]. These costs were based on subjects with principal diagnoses of pneumococcal meningitis, all-cause bacteremia, and pneumococcal pneumonia, and were available by age group. Costs, including the 6.581% increase for physician costs [119], were $20,401 - $26,933 for meningitis, $9342 - $13,311 for bacteremia, and $11,439 - $13,006 for pneumonia. These figures resulted in an estimated 7.8 billion dollars in prevented direct medical costs through 2010, and an additional 19.9 billion dollars in estimated savings through 2025.

The differences in cost estimates between the Duke hospital system and HCUP data led to significant differences in estimates of total medical costs, reductions in costs, and cost savings. These differences should not be taken to indicate that the Duke hospital system has higher costs than the average hospital, as two different types of costs were being compared.

Costs from the HCUP data were limited to subjects with a particular primary diagnosis, based on an ICD-9-CM code. Other research has indicated that primary diagnoses may be lacking in specificity (such as a diagnosis of code 486.0, pneumonia, organism unspecified) or misdiagnosed [138]. A comparison of the numbers of patients with a primary diagnosis of pneumonia (Clinical Classification Software [CCS] code 122) against the numbers of patients with pneumonia listed as either a primary or secondary diagnosis during 1999 – 2009

indicated that the fraction of subjects for whom pneumonia was the primary diagnosis declined from 62% to 42%. Therefore, many subjects were not included in these average costs, since costs were only available based on primary diagnosis.

Furthermore, due to the lack of a specific ICD-9-CM code for pneumococcal bacteremia, costs for all-cause bacteremia were used in the supplemental analysis. The CDC ABCs characterized subjects with isolates obtained from the blood as bacteremia (i.e., the presence of bacteria in the blood). These cases were not differentiated from septicemia, defined as a serious infection of the blood. Hospitalization costs of subjects with a primary diagnosis of pneumococcal septicemia ranged from $20,607 to $25,350 by age group, while similar estimates of all-cause bacteremia ranged from $9,342 to $13,331. Since the CDC data did not state which cases of bacteremia were actually septicemia, we used costs for bacteremia in our calculations. This likely led to an underestimation of the actual costs of pneumococcal blood infections.

In contrast, the Duke hospital system costs were the average hospitalization cost for subjects with a pneumococcal isolate from the lower respiratory tract or a normally sterile site, regardless of diagnoses. Therefore, these costs were not limited to subjects with a diagnosis of pneumococcal disease. This had the advantage of including costs for subjects with diagnoses of respiratory distress or unspecified pneumonia, whose microbiological laboratory results indicated pneumococcal pneumonia, or subjects who had a comorbidity listed as a primary diagnosis. In all cases, patients included in this sample had sufficient symptoms to cause medical staff to collect a sample from the appropriate site and send it to the microbiological laboratory for identification. However, these figures may also include unrelated medical costs.

In summary, the HCUP data was limited to a strict subset of patients, based on primary diagnosis, and very likely did not include all patients with pneumococcal disease. In addition, bacteremia costs may have been underestimated. The costs from the Duke hospital

system encompassed a wider range of subjects with pneumococcal disease, and were likely to include costs for conditions unrelated to the disease. The true cost of the disease was likely somewhere between the two. Therefore, while each type of cost calculation has its advantages and disadvantages, together they provide a range of possible actual costs.

Regardless of which cost estimates were used, we concluded that the PCV7 vaccine was still cost-effective. In the primary analysis, the vaccine was estimated to save 63.7 billion dollars during the first ten years, and an additional 151.2 billion dollars by the end of 2025. The supplemental analysis, utilizing the HCUP data, estimated lower cost savings (7.8 billion dollars through 2009, and an additional 19.9 billion dollars by the end of 2025).

This was a very different conclusion than was reached in some earlier cost-effectiveness analyses. The first cost-effectiveness analysis of PCV7 in the United States, evaluating a hypothetical birth cohort of 3.8 million subjects, concluded that the vaccine would not be cost-effective in terms of medical costs if the vaccine cost more than $18 per dose [62]. A later cost-effectiveness study in the United Kingdom concluded that the vaccine would be cost-effective at half of its list price [63]. A more recent cost-effectiveness study in the United States, which incorporated indirect vaccine effects, concluded that the vaccine was cost-effective [64].

Challenges

S. pneumoniae is a challenging epidemiological system to characterize and understand, both from a data collection perspective, and from a modeling standpoint. The majority of transmission occurs between healthy, asymptomatically colonized subjects, making transmission difficult to identify, and subsequent colonization difficult to quantify. While IPD is reported by subjects who were treated at acute care hospitals, and subsequently have isolates sent to a microbiological laboratory, colonization information must be collected by conducting a survey. In such surveys, nasopharyngeal swabs are used to collect samples from subjects repeatedly over time. Weekly swabs over a six-month period might be

sufficient to quantify the frequency and average duration of colonization among young children, since children under two years old have a 10 – 16 week estimated average duration of colonization, 14 week time to reinfection [82], and approximately 50% of those children are colonized at any time. A significantly longer survey period or larger sample size would be needed to accurately characterize colonization among adults, who have a 4 – 7 week estimated average duration of colonization and approximately 8% are colonized at any time. Each swab would need to be cultured, to identify any bacteria, and the pneumococcal bacteria serotyped. This type of survey could be used to obtain estimates for average durations of colonization by serotype group and age group. Such a survey could also record the average time to recolonization, in order to better estimate the average duration of temporary immunity.

Although a limited number of such surveys have been conducted in recent decades, such information is expensive to collect. That information is then proprietary, with publicly available data often limited to information given in published literature.

Several parameter values, such as transmission probabilities (p_i) and average duration of temporary immunity (ρ_j^{-1}), are almost impossible to directly estimate in a clinical setting, and have not previously been explicitly estimated using models. We made assumptions regarding the relationships between transmission probabilities among age groups and serotype groups, in order to reduce eighteen unique parameter values down to four unique values (p_1, p_2, p_3, d). Similarly, the average duration of temporary immunity was taken to be the same among subjects colonized with all serotypes ($\rho_1 = \rho_2$), in order to reduce the total number of unknown parameter values for the colonization model to five, and to fit those five parameter values using five known outcomes. We acknowledge that this is not an ideal statistical approach, and that it would be improved with more data; but, as with many models of this type, we were constrained by the amount of information available. Detailed results from a longitudinal study of colonization would be necessary to better characterize these

parameter values. In spite of the shortcomings of this approach, such as the overfitting of the model, this method allowed us to use data-driven estimates for these parameter values instead of choosing values without a numeric basis.

A similar optimization approach was then used to estimate the fractions of colonized subjects developing IPD, with total IPD incidence for each of eleven years being used to fit seven parameter values: the effect of vaccination on VT colonization (e) and six case:colonization ratios (f_{ij}). While, again, this approach is not statistically optimal, it allows us to use the data available, instead of selecting values without numerical basis. Other parameter values used in the model were based on clinical data in the published literature, but were often based on data from a combination of studies, due to the limited amount of such data available.

Another simplification used in this model was the method of grouping multiple serotypes together into the VT and NVT serotype groups, and treating each of these groups as a single serotype, as has been done in virtually every other pneumococcal transmission model [44; 46; 47]. This assumed that transmission probabilities and average durations of colonization are the same within each age group and serotype group, and that a subject could not concurrently be colonized with more than one VT serotype, or with more than one NVT serotype. We considered this a necessary simplification, due to the lack of information available on specific serotypes.

Comprehensive information on pneumococcal disease was also difficult to obtain. The CDC ABCs team reports IPD estimates by age group and isolate location (meningitis; bacteremia without focus; pneumonia with bacteremia), based on patients who were treated at acute care hospitals within certain surveillance areas, and have isolates sent to a microbiological laboratory. These reported IPD estimates do not include subjects who were not treated at acute care hospitals, or who did not have isolates sent to a laboratory for identification. Non-invasive pneumococcal pneumonia is also excluded from these disease estimates. The CDC does collect serotype information on these isolates, but does not provide publicly available

information on the amount of reported IPD caused by VT or NVT serotypes. Therefore, we fitted the model-based estimates of VT and NVT colonization to total IPD incidence to determine case:colonization ratios for each serotype group, and the fractions of IPD caused by VT serotypes within age group in 1999 and 2007 [73], since we were unable to directly determine the incidence of VT IPD and NVT IPD each year.

Pneumococcal pneumonia is poorly characterized in the literature. Determination of the microbiological cause is not necessary for a diagnosis of pneumonia. In a 2009 nationwide inpatient sample, 84% of diagnosed pneumonia was classified as being due to an unspecified organism or unspecified bacterium (ICD-9-CM codes of 482.9, 485, 486) [96]. During 1999-2009, *S. pneumoniae* accounted for an average of 15.4% of pneumonia diagnoses that were not due to unspecified organisms or unspecified bacteria [96]. We then estimated the 1999 pneumococcal pneumonia incidence as 15.4% of all pneumonia diagnoses. Using the ten-year average likely underestimated the fraction of pneumonia caused by *S. pneumoniae*, which led to a conservative estimate of pneumonia incidence, since *S. pneumoniae* accounted for 18.2% of those pneumonia diagnoses in 1999, and other research has found it to cause approximately 30% of pneumonia cases [94; 98].

We acknowledge that the model-based estimates of pneumococcal pneumonia incidence were very sensitive to slight changes to parameter values used in their calculation. However, the estimated cost savings were so large that the pneumococcal vaccine would still be cost-effective if reductions in pneumonia costs were only 15% of those we estimated. In other words, we would still consider the vaccine to be cost effective across very different characterizations of pneumonia effects. Therefore, although these calculations of pneumococcal pneumonia costs were based on a poorly characterized process, the conclusion that the vaccine is cost-effective was very robust to differences in pneumonia characterization.

Huang et al. [136] evaluated the estimated pneumococcal disease burden in 2004, and included other types of pneumococcal infection, such as acute otitis media, sinusitis, and bone/joint infection, and included indirect costs such as work-loss, lost wages, and long-term disability costs. These were beyond the scope of our current research. Our analysis focused on the types of disease leading to the greatest medical costs (i.e., those which were most prevalent and which were primarily treated on an inpatient basis). We acknowledge that these cost estimates did not include all possible types of illness caused by *S. pneumoniae*. However, we took the incidence of all types of pneumococcal disease to be proportional to colonization. Therefore, the decreases in colonization due to vaccination would also cause decreases in these other types of pneumococcal disease. For this reason, we considered our estimates of cost savings due to prevented pneumococcal disease to be conservative underestimates.

Further Research

As previously mentioned, an expanded 13-serotype vaccine, PCV13, was approved for use in the United States in February 2010. This vaccine includes the seven serotypes in the original PCV7 vaccine, as well as serotypes 1, 3, 5, 6A, 7F, and 19A.

The next step in this research would be to expand the existing model to incorporate three groups of serotypes: those included in the PCV7 vaccine, those included in the PCV13 vaccine but not the PCV7 vaccine, and those not included in either vaccine. The existing model could then be used to estimate the colonization and IPD incidence at the time of the PCV13 introduction. Those incidence levels would be used as the baseline values in the expanded model, which would track PCV7, PCV13, and NVT serotypes separately. That model would then be able to predict the additional declines in IPD and colonization caused by the expanded vaccine, and be used to evaluate the cost-effectiveness of the new vaccine.

The current model could also be expanded to evaluate the effect of lumping all serotypes into groups. The 3 colonization statuses currently used for each serotype (S, I, R) could be

expanded to include two additional types: dually colonized by two serotypes within the same serotype group, and colonized by one serotype and recovered from the other serotype within the same group. Results from that model could be compared to the current model to evaluate the impact of the assumption that only one serotype from each group can be present within a subject at a certain time.

REFERENCES

1. World Health Organization. Immunization surveillance, assessment and monitoring. 2011 Oct 19]. Available from: http://www.who.int/immunization_monitoring/burden/Pneumo_hib_estimates/en/index2.html.

2. Bogaert D, De Groot R, Hermans PW. 2004. Streptococcus pneumoniae colonisation: the key to pneumococcal disease. Lancet Infect Dis 4(3):144-54.

3. Kayhty H, Auranen K, Nohynek H, Dagan R, Makela H. 2006. Nasopharyngeal colonization: a target for pneumococcal vaccination. Expert Rev Vaccines 5(5):651-67.

4. McCool TL, Cate TR, Moy G, Weiser JN. 2002. The immune response to pneumococcal proteins during experimental human carriage. J Exp Med 195(3):359-65.

5. Goldblatt D, Hussain M, Andrews N, Ashton L, Virta C, Melegaro A, Pebody R, George R, Soininen A, Edmunds J, Gay N, Kayhty H, Miller E. 2005. Antibody responses to nasopharyngeal carriage of Streptococcus pneumoniae in adults: a longitudinal household study. J Infect Dis 192(3):387-93.

6. Hansman D, Morris S, Gregory M, McDonald B. 1985. Pneumococcal carriage amongst Australian aborigines in Alice Springs, Northern Territory. J Hyg (Lond) 95(3):677-84.

7. Melegaro A, Gay NJ, Medley GF. 2004. Estimating the transmission parameters of pneumococcal carriage in households. Epidemiol Infect 132(3):433-41.

8. Regev-Yochay G, Raz M, Dagan R, Porat N, Shainberg B, Pinco E, Keller N, Rubinstein E. 2004. Nasopharyngeal carriage of Streptococcus pneumoniae by adults and children in community and family settings. Clin Infect Dis 38(5):632-9.

9. Ekdahl K, Ahlinder I, Hansson HB, Melander E, Molstad S, Soderstrom M, Persson K. 1997. Duration of nasopharyngeal carriage of penicillin-resistant Streptococcus pneumoniae: experiences from the South Swedish Pneumococcal Intervention Project. Clin Infect Dis 25(5):1113-7.

10. Auranen K, Arjas E, Leino TM, Takala A. 2000. Transmission of Pneumococcal Carriage in Families: A Latent Markov Process Model for Binary Longitudinal Data. J Am Stat Assoc 95(452):1044-1053.

11. Ekholm A, Jokinen J, Kilpi T. 2002. Combining regression and association modelling for longitudinal data on bacterial carriage. Stat Med 21(5):773-91.

12. Meats E, Brueggemann AB, Enright MC, Sleeman K, Griffiths DT, Crook DW, Spratt BG. 2003. Stability of serotypes during nasopharyngeal carriage of Streptococcus pneumoniae. J Clin Microbiol 41(1):386-92.

13. Gray BM, Converse GM, 3rd, Dillon HC, Jr. 1980. Epidemiologic studies of Streptococcus pneumoniae in infants: acquisition, carriage, and infection during the first 24 months of life. J Infect Dis 142(6):923-33.

14. Centers for Disease Control and Prevention. 2011. Epidemiology and Prevention of Vaccine-Preventable Diseases. In: Atkinson W, Wolfe S, Hamborsky J, editors. 12 ed.: Washington DC: Public Health Foundation.

15. Centers for Disease Control and Prevention. 1997. Prevention of pneumococcal disease: recommendations of the Advisory Committee on Immunization Practices (ACIP). MMWR Recomm Rep 46(RR-8):1-24.

16. Sutton KL, Banks HT, Castillo-Chavez C. 2008. Estimation of invasive pneumococcal disease dynamics parameters and the impact of conjugate vaccination in Australia. Math Biosci Eng 5(1):175-204.

17. Centers for Disease Control and Prevention. Active Bacterial Core Surveillance Report, Emerging Infections Program Network, *Streptococcus pneumoniae*. 2011 Oct 19]. Available from: http://www.cdc.gov/abcs/reports-findings/surv-reports.html

18. De Graeve D, Beutels P. 2004. Economic aspects of pneumococcal pneumonia: a review of the literature. Pharmacoeconomics 22(11):719-40.

19. Centers for Disease Control and Prevention. Update: pneumococcal polysaccharide vaccine usage--United States. Recommendations of the Immunization Practices Advisory Committee. Morbidity and Mortality Weekly Report [Internet]. 2011 Oct 19];33(20):273-6, 281. Available from: http://www.cdc.gov/mmwr/preview/mmwrhtml/00031107.htm

20. World Health Organization. Pneumococcal vaccines. WHO position paper. Wkly Epidemiol Rec [Internet]. 2011 Oct 19];83:373-384. Available from: http://www.who.int/wer/2008/wer8342.pdf

21. Black S, Shinefield H, Baxter R, Austrian R, Bracken L, Hansen J, Lewis E, Fireman B. 2004. Postlicensure surveillance for pneumococcal invasive disease after use of heptavalent pneumococcal conjugate vaccine in Northern California Kaiser Permanente. Pediatr Infect Dis J 23(6):485-9.

22. Whitney CG, Farley MM, Hadler J, Harrison LH, Bennett NM, Lynfield R, Reingold A, Cieslak PR, Pilishvili T, Jackson D, Facklam RR, Jorgensen JH, Schuchat A. 2003. Decline in invasive pneumococcal disease after the introduction of protein-polysaccharide conjugate vaccine. N Engl J Med 348(18):1737-46.

23. Hansen J, Black S, Shinefield H, Cherian T, Benson J, Fireman B, Lewis E, Ray P, Lee J. 2006. Effectiveness of heptavalent pneumococcal conjugate vaccine in children younger than 5 years of age for prevention of pneumonia: updated analysis using World Health Organization standardized interpretation of chest radiographs. Pediatr Infect Dis J 25(9):779-81.

24. Lexau CA, Lynfield R, Danila R, Pilishvili T, Facklam R, Farley MM, Harrison LH, Schaffner W, Reingold A, Bennett NM, Hadler J, Cieslak PR, Whitney CG. 2005. Changing epidemiology of invasive pneumococcal disease among older adults in the era of pediatric pneumococcal conjugate vaccine. JAMA 294(16):2043-51.

25. Haddy RI, Perry K, Chacko CE, Helton WB, Bowling MG, Looney SW, Buck GE. 2005. Comparison of incidence of invasive Streptococcus pneumoniae disease among children before and after introduction of conjugated pneumococcal vaccine. Pediatr Infect Dis J 24(4):320-3.

26. Byington CL, Samore MH, Stoddard GJ, Barlow S, Daly J, Korgenski K, Firth S, Glover D, Jensen J, Mason EO, Shutt CK, Pavia AT. 2005. Temporal trends of invasive disease due to Streptococcus pneumoniae among children in the intermountain west: emergence of nonvaccine serogroups. Clin Infect Dis 41(1):21-9.

27. Hsu K, Pelton S, Karumuri S, Heisey-Grove D, Klein J. 2005. Population-based surveillance for childhood invasive pneumococcal disease in the era of conjugate vaccine. Pediatr Infect Dis J 24(1):17-23.

28. Kaplan SL, Mason EO, Jr., Wald ER, Schutze GE, Bradley JS, Tan TQ, Hoffman JA, Givner LB, Yogev R, Barson WJ. 2004. Decrease of invasive pneumococcal

infections in children among 8 children's hospitals in the United States after the introduction of the 7-valent pneumococcal conjugate vaccine. Pediatrics 113(3 Pt 1):443-9.

29. Poehling KA, Talbot TR, Griffin MR, Craig AS, Whitney CG, Zell E, Lexau CA, Thomas AR, Harrison LH, Reingold AL, Hadler JL, Farley MM, Anderson BJ, Schaffner W. 2006. Invasive pneumococcal disease among infants before and after introduction of pneumococcal conjugate vaccine. JAMA 295(14):1668-74.

30. Stephens DS, Zughaier SM, Whitney CG, Baughman WS, Barker L, Gay K, Jackson D, Orenstein WA, Arnold K, Schuchat A, Farley MM. 2005. Incidence of macrolide resistance in Streptococcus pneumoniae after introduction of the pneumococcal conjugate vaccine: population-based assessment. Lancet 365(9462):855-63.

31. Cohen R, Levy C, Bonnet E, Grondin S, Desvignes V, Lecuyer A, Fritzell B, Varon E. 2010. Dynamic of pneumococcal nasopharyngeal carriage in children with acute otitis media following PCV7 introduction in France. Vaccine 28(37):6114-21.

32. Dagan R, Givon-Lavi N, Zamir O, Sikuler-Cohen M, Guy L, Janco J, Yagupsky P, Fraser D. 2002. Reduction of nasopharyngeal carriage of Streptococcus pneumoniae after administration of a 9-valent pneumococcal conjugate vaccine to toddlers attending day care centers. J Infect Dis 185(7):927-36.

33. Grivea IN, Tsantouli AG, Michoula AN, Syrogiannopoulos GA. 2011. Dynamics of Streptococcus pneumoniae nasopharyngeal carriage with high heptavalent pneumococcal conjugate vaccine coverage in Central Greece. Vaccine 29(48):8882-7.

34. Huang SS, Hinrichsen VL, Stevenson AE, Rifas-Shiman SL, Kleinman K, Pelton SI, Lipsitch M, Hanage WP, Lee GM, Finkelstein JA. 2009. Continued impact of pneumococcal conjugate vaccine on carriage in young children. Pediatrics 124(1):e1-11.

35. Mbelle N, Huebner RE, Wasas AD, Kimura A, Chang I, Klugman KP. 1999. Immunogenicity and impact on nasopharyngeal carriage of a nonavalent pneumococcal conjugate vaccine. J Infect Dis 180(4):1171-6.

36. Rodrigues F, Nunes S, Sa-Leao R, Goncalves G, Lemos L, de Lencastre H. 2009. Streptococcus pneumoniae nasopharyngeal carriage in children attending day-care centers in the central region of Portugal, in the era of 7-valent pneumococcal conjugate vaccine. Microb Drug Resist 15(4):269-77.

37. Vestrheim DF, Hoiby EA, Aaberge IS, Caugant DA. 2010. Impact of a pneumococcal conjugate vaccination program on carriage among children in Norway. Clin Vaccine Immunol 17(3):325-34.

38. Millar EV, O'Brien KL, Watt JP, Bronsdon MA, Dallas J, Whitney CG, Reid R, Santosham M. 2006. Effect of community-wide conjugate pneumococcal vaccine use in infancy on nasopharyngeal carriage through 3 years of age: a cross-sectional study in a high-risk population. Clin Infect Dis 43(1):8-15.

39. O'Brien KL, Millar EV, Zell ER, Bronsdon M, Weatherholtz R, Reid R, Becenti J, Kvamme S, Whitney CG, Santosham M. 2007. Effect of pneumococcal conjugate vaccine on nasopharyngeal colonization among immunized and unimmunized children in a community-randomized trial. J Infect Dis 196(8):1211-20.

40. Vestrheim DF, Hoiby EA, Aaberge IS, Caugant DA. 2008. Phenotypic and genotypic characterization of Streptococcus pneumoniae strains colonizing children attending day-care centers in Norway. J Clin Microbiol 46(8):2508-18.

41. Centers for Disease Control and Prevention. 2010. Licensure of a 13-valent pneumococcal conjugate vaccine (PCV13) and recommendations for use among children - Advisory Committee on Immunization Practices (ACIP). MMWR Morb Mortal Wkly Rep 59(9):258-61.

42. Lipsitch M. 1997. Vaccination against colonizing bacteria with multiple serotypes. Proc Natl Acad Sci U S A 94(12):6571-6.

43. Zhang Y, Auranen K, Eichner M. 2004. The influence of competition and vaccination on the coexistence of two pneumococcal serotypes. Epidemiol Infect 132(6):1073-81.

44. Temime L, Guillemot D, Boelle PY. 2004. Short- and long-term effects of pneumococcal conjugate vaccination of children on penicillin resistance. Antimicrob Agents Chemother 48(6):2206-13.

45. Edmunds WJ, O'Callaghan CJ, Nokes DJ. 1997. Who mixes with whom? A method to determine the contact patterns of adults that may lead to the spread of airborne infections. Proc Biol Sci 264(1384):949-57.

46. Melegaro A, Choi YH, George R, Edmunds WJ, Miller E, Gay NJ. 2010. Dynamic models of pneumococcal carriage and the impact of the Heptavalent Pneumococcal Conjugate Vaccine on invasive pneumococcal disease. BMC Infect Dis 10:90.

47. Snedecor SJ, Strutton DR, Ciuryla V, Schwartz EJ, Botteman MF. 2009. Transmission-dynamic model to capture the indirect effects of infant vaccination with Prevnar (7-valent pneumococcal conjugate vaccine (PCV7)) in older populations. Vaccine 27(34):4694-703.

48. Van Effelterre T, Moore MR, Fierens F, Whitney CG, White L, Pelton SI, Hausdorff WP. 2010. A dynamic model of pneumococcal infection in the United States: implications for prevention through vaccination. Vaccine 28(21):3650-60.

49. Fine MJ, Smith MA, Carson CA, Mutha SS, Sankey SS, Weissfeld LA, Kapoor WN. 1996. Prognosis and outcomes of patients with community-acquired pneumonia. A meta-analysis. JAMA 275(2):134-41.

50. Schuchat A, Robinson K, Wenger JD, Harrison LH, Farley M, Reingold AL, Lefkowitz L, Perkins BA. 1997. Bacterial meningitis in the United States in 1995. Active Surveillance Team. N Engl J Med 337(14):970-6.

51. Birnbaum HG, Morley M, Greenberg PE, Cifaldi M, Colice GL. 2001. Economic burden of pneumonia in an employed population. Arch Intern Med 161(22):2725-31.

52. Diekema DJ, Brueggemann AB, Doern GV. 2000. Antimicrobial-drug use and changes in resistance in Streptococcus pneumoniae. Emerg Infect Dis 6(5):552-6.

53. Abb J. 2005. Ten-year surveillance of antimicrobial resistance of Streptococcus pneumoniae in a south-west German teaching hospital. Int J Infect Dis 9(1):60-1.

54. Castillo EM, Rickman LS, Brodine SK, Ledbetter EK, Kelly C. 2000. Streptococcus pneumoniae: bacteremia in an era of penicillin resistance. Am J Infect Control 28(3):239-43.

55. Jacobs MR, Good CE, Beall B, Bajaksouzian S, Windau AR, Whitney CG. 2008. Changes in serotypes and antimicrobial susceptibility of invasive Streptococcus pneumoniae strains in Cleveland: a quarter century of experience. J Clin Microbiol 46(3):982-90.

56. Reacher MH, Shah A, Livermore DM, Wale MC, Graham C, Johnson AP, Heine H, Monnickendam MA, Barker KF, James D, George RC. 2000. Bacteraemia and antibiotic resistance of its pathogens reported in England and Wales between 1990 and 1998: trend analysis. BMJ 320(7229):213-6.

57. Fiore AE, Moroney JF, Farley MM, Harrison LH, Patterson JE, Jorgensen JH, Cetron M, Kolczak MS, Breiman RF, Schuchat A. 2000. Clinical outcomes of meningitis caused by Streptococcus pneumoniae in the era of antibiotic resistance. Clin Infect Dis 30(1):71-7.

58. Cosgrove SE, Qi Y, Kaye KS, Harbarth S, Karchmer AW, Carmeli Y. 2005. The impact of methicillin resistance in Staphylococcus aureus bacteremia on patient outcomes: mortality, length of stay, and hospital charges. Infect Control Hosp Epidemiol 26(2):166-74.

59. Cosgrove SE, Carmeli Y. 2003. The impact of antimicrobial resistance on health and economic outcomes. Clin Infect Dis 36(11):1433-7.

60. Reed SD, Blough DK, Meyer K, Jarvik JG. 2001. Inpatient costs, length of stay, and mortality for cerebrovascular events in community hospitals. Neurol 57(2):305-14.

61. Centers for Disease Control and Prevention. CDC vaccine price list. 2011 Jul 22]. Available from: http://www.cdc.gov/vaccines/programs/vfc/cdc-vac-price-list-archives.htm

62. Lieu TA, Ray GT, Black SB, Butler JC, Klein JO, Breiman RF, Miller MA, Shinefield HR. 2000. Projected cost-effectiveness of pneumococcal conjugate vaccination of healthy infants and young children. JAMA 283(11):1460-8.

63. Melegaro A, Edmunds WJ. 2004. Cost-effectiveness analysis of pneumococcal conjugate vaccination in England and Wales. Vaccine 22(31-32):4203-14.

64. Ray GT, Pelton SI, Klugman KP, Strutton DR, Moore MR. 2009. Cost-effectiveness of pneumococcal conjugate vaccine: an update after 7 years of use in the United States. Vaccine 27(47):6483-94.

65. Obaro SK, Adegbola RA, Banya WA, Greenwood BM. 1996. Carriage of pneumococci after pneumococcal vaccination. Lancet 348(9022):271-2.

66. Hammitt LL, Bruden DL, Butler JC, Baggett HC, Hurlburt DA, Reasonover A, Hennessy TW. 2006. Indirect Effect of Conjugate Vaccine on Adult Carriage of Streptococcus pneumoniae: An Explanation of Trends in Invasive Pneumococcal Disease. J Infect Dis 193(11):1487-94.

67. Ghaffar F, Barton T, Lozano J, Muniz LS, Hicks P, Gan V, Ahmad N, McCracken GH, Jr. 2004. Effect of the 7-valent pneumococcal conjugate vaccine on

nasopharyngeal colonization by Streptococcus pneumoniae in the first 2 years of life. Clin Infect Dis 39(7):930-8.

68. Whitney CG, Pilishvili T, Farley MM, Schaffner W, Craig AS, Lynfield R, Nyquist AC, Gershman KA, Vazquez M, Bennett NM, Reingold A, Thomas A, Glode MP, Zell ER, Jorgensen JH, Beall B, Schuchat A. 2006. Effectiveness of seven-valent pneumococcal conjugate vaccine against invasive pneumococcal disease: a matched case-control study. Lancet 368(9546):1495-502.

69. Black S, Shinefield H, Fireman B, Lewis E, Ray P, Hansen JR, Elvin L, Ensor KM, Hackell J, Siber G, Malinoski F, Madore D, Chang I, Kohberger R, Watson W, Austrian R, Edwards K. 2000. Efficacy, safety and immunogenicity of heptavalent pneumococcal conjugate vaccine in children. Northern California Kaiser Permanente Vaccine Study Center Group. Pediatr Infect Dis J 19(3):187-95.

70. Erasto P, Hoti F, Granat SM, Mia Z, Makela PH, Auranen K. 2010. Modelling multi-type transmission of pneumococcal carriage in Bangladeshi families. Epidemiol Infect 138(6):861-72.

71. Choi YH, Jit M, Gay N, Andrews N, Waight PA, Melegaro A, George R, Miller E. 2011. 7-valent pneumococcal conjugate vaccination in England and wales: is it still beneficial despite high levels of serotype replacement? PLoS One 6(10):e26190.

72. Robinson KA, Baughman W, Rothrock G, Barrett NL, Pass M, Lexau C, Damaske B, Stefonek K, Barnes B, Patterson J, Zell ER, Schuchat A, Whitney CG. 2001. Epidemiology of invasive Streptococcus pneumoniae infections in the United States, 1995-1998: Opportunities for prevention in the conjugate vaccine era. JAMA 285(13):1729-35.

73. Pilishvili T, Lexau C, Farley MM, Hadler J, Harrison LH, Bennett NM, Reingold A, Thomas A, Schaffner W, Craig AS, Smith PJ, Beall BW, Whitney CG, Moore MR. 2010. Sustained reductions in invasive pneumococcal disease in the era of conjugate vaccine. J Infect Dis 201(1):32-41.

74. Malley R, Trzcinski K, Srivastava A, Thompson CM, Anderson PW, Lipsitch M. 2005. CD4+ T cells mediate antibody-independent acquired immunity to pneumococcal colonization. Proc Natl Acad Sci U S A 102(13):4848-53.

75. Weinberger DM, Dagan R, Givon-Lavi N, Regev-Yochay G, Malley R, Lipsitch M. 2008. Epidemiologic evidence for serotype-specific acquired immunity to pneumococcal carriage. J Infect Dis 197(11):1511-8.

76. Gwaltney JM, Jr., Sande MA, Austrian R, Hendley JO. 1975. Spread of Streptococcus pneumoniae in families. II. Relation of transfer of S. pneumoniae to incidence of colds and serum antibody. J Infect Dis 132(1):62-8.

77. Hussain M, Melegaro A, Pebody RG, George R, Edmunds WJ, Talukdar R, Martin SA, Efstratiou A, Miller E. 2005. A longitudinal household study of Streptococcus pneumoniae nasopharyngeal carriage in a UK setting. Epidemiol Infect 133(5):891-8.

78. Heron MP, Hoyert DL, Murphy SL, Xu JQ, Kochanek KD, Betzaida T-V. 2009. Deaths: Final data for 2006. National Vital Statistics Reports. Hyattsville, MD: National Center for Health Statistics.

79. Mossong J, Hens N, Jit M, Beutels P, Auranen K, Mikolajczyk R, Massari M, Salmaso S, Tomba GS, Wallinga J, Heijne J, Sadkowska-Todys M, Rosinska M, Edmunds WJ. 2008. Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS Med 5(3):e74.

80. Anderson RM, May RM. 1985. Age-related changes in the rate of disease transmission: implications for the design of vaccination programmes. J Hyg (Lond) 94(3):365-436.

81. Sleeman KL, Griffiths D, Shackley F, Diggle L, Gupta S, Maiden MC, Moxon ER, Crook DW, Peto TE. 2006. Capsular serotype-specific attack rates and duration of carriage of Streptococcus pneumoniae in a population of children. J Infect Dis 194(5):682-8.

82. Raymond J, Le Thomas I, Moulin F, Commeau A, Gendrel D, Berche P. 2000. Sequential colonization by Streptococcus pneumoniae of healthy children living in an orphanage. J Infect Dis 181(6):1983-8.

83. Lagarias JC, Reeds J.A., Wright M. H., Wright P.E. 1998. Convergence properties of the Nelder-Mead simplex method in low dimensions. SIAM J Optim 9(1):112-147.

84. U.S. Department of Health and Human Services (DHHS). National Center for Health Statistics. National Immunization Surveys, 2001 - 2009 [Internet]. Hyattsville, MD: Centers for Disease Control and Prevention. Available from: http://www.cdc.gov/nchs/nis.htm

85. Nuorti JP, Martin SW, Smith PJ, Moran JS, Schwartz B. 2008. Uptake of pneumococcal conjugate vaccine among children in the 1998-2002 United States birth cohorts. Am J Prev Med 34(1):46-53.

86. Martin JA, Hamilton BE, Sutton PD, Ventura SJ, Mathews T, Kirmeyer S, Osterman MJ. 2010. Births: Final data for 2007. National Vital Statistics Reports. Hyattsville, MD: National Center for Health Statistics.

87. Diekmann O, Heesterbeek JAP. 2000. Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation. Wiley.

88. Blower SM, Hartel D, Dowlatabadi H, Anderson RM, May RM. 1991. Drugs, sex and HIV: a mathematical model for New York City. Philos Trans R Soc Lond B Biol Sci 331(1260):171-87.

89. Marino S, Hogue IB, Ray CJ, Kirschner DE. 2008. A methodology for performing global uncertainty and sensitivity analysis in systems biology. J Theor Biol 254(1):178-96.

90. Blower SM, Dowlatabadi H. 1994. Sensitivity and Uncertainty Analysis of Complex Models of Disease Transmission: An HIV Model, as an Example. Int Stat Rvw 62(2):229-243.

91. Brugger SD, Frey P, Aebi S, Hinds J, Muhlemann K. 2010. Multiple colonization with S. pneumoniae before and after introduction of the seven-valent conjugated pneumococcal polysaccharide vaccine. PLoS One 5(7):e11638.

92. Lipsitch M. 2001. Interpreting results from trials of pneumococcal conjugate vaccines: a statistical test for detecting vaccine-induced increases in carriage of nonvaccine serotypes. Am J Epidemiol 154(1):85-92.

93. Turner P, Hinds J, Turner C, Jankhot A, Gould K, Bentley SD, Nosten F, Goldblatt D. 2011. Improved Detection of Nasopharyngeal Cocolonization by Multiple Pneumococcal Serotypes by Use of Latex Agglutination or Molecular Serotyping by Microarray. J Clin Microbiol 49(5):1784-1789.

94. Fine MJ, Stone RA, Singer DE, Coley CM, Marrie TJ, Lave JR, Hough LJ, Obrosky DS, Schulz R, Ricci EM, Rogers JC, Kapoor WN. 1999. Processes and outcomes of care for patients with community-acquired pneumonia: results from the Pneumonia Patient Outcomes Research Team (PORT) cohort study. Arch Intern Med 159(9):970-80.

95. Xu JQ, Kochanek KD, Murphy SL, Betzaida T-V. 2010. Deaths: Final data for 2007. National Vital Statistics Reports. Hyattsville, MD: National Center for Health Statistics.

96. Agency for Healthcare Research and Quality, US Dept. of Health and Human Services. Healthcare Cost and Utilization Project (HCUP). 2011 Oct 19]. Available from: http://hcupnet.ahrq.gov/

97. Brown PD, Lerner SA. 1998. Community-acquired pneumonia. Lancet 352(9136):1295-302.

98. Ruiz-Gonzalez A, Falguera M, Nogues A, Rubio-Caballero M. 1999. Is Streptococcus pneumoniae the leading cause of pneumonia of unknown etiology? A microbiologic study of lung aspirates in consecutive patients with community-acquired pneumonia. Am J Med 106(4):385-90.

99. Gray BM, Converse GM, 3rd, Dillon HC, Jr. 1979. Serotypes of Streptococcus pneumoniae causing disease. J Infect Dis 140(6):979-83.

100. Manning WG, Mullahy J. 2001. Estimating log models: to transform or not to transform? J Health Econ 20(4):461-94.

101. Duan N. 1983. Smearing estimate: a nonparametric retransformation method. J Amer Stat Assn 78:605-610.

102. US Bureau of Labor Statistics. Consumer Price Index - All Urban Consumers (CPI-U), 10-19-2011. 2011 Oct 19]. Available from: ftp://ftp.bls.gov/pub/special.requests/cpi/cpiai.txt

103. Kastenbauer S, Pfister HW. 2003. Pneumococcal meningitis in adults: spectrum of complications and prognostic factors in a series of 87 cases. Brain 126(Pt 5):1015-25.

104. Weisfelt M, van de Beek D, Spanjaard L, Reitsma JB, de Gans J. 2006. Clinical features, complications, and outcome in adults with pneumococcal meningitis: a prospective case series. Lancet Neurol 5(2):123-9.

105. McKenzie H, Reid N, Dijkhuizen RS. 2000. Clinical and microbiological epidemiology of Streptococcus pneumoniae bacteraemia. J Med Microbiol 49(4):361-6.

106. Mufson MA, Stanek RJ. 1999. Bacteremic pneumococcal pneumonia in one American City: a 20-year longitudinal study, 1978-1997. Am J Med 107(1A):34S-43S.

107. Plouffe JF, Breiman RF, Facklam RR. 1996. Bacteremia with Streptococcus pneumoniae. Implications for therapy and prevention. Franklin County Pneumonia Study Group. JAMA 275(3):194-8.

108. Raz R, Elhanan G, Shimoni Z, Kitzes R, Rudnicki C, Igra Y, Yinnon A. 1997. Pneumococcal bacteremia in hospitalized Israeli adults: epidemiology and resistance to penicillin. Israeli Adult Pneumococcal Bacteremia Group. Clin Infect Dis 24(6):1164-8.

109. Chi RC, Jackson LA, Neuzil KM. 2006. Characteristics and outcomes of older adults with community-acquired pneumococcal bacteremia. J Am Geriatr Soc 54(1):115-20.

110. Altclas J, Salgueira C, DiMartino A. 2004. Pneumococcal bacteremia in a single center in Argentina. Int J Infect Dis 8(1):53-58.

111. Song JH, Jung SI, Ki HK, Shin MH, Ko KS, Son JS, Chang HH, Kim SW, Lee H, Kim YS, Oh WS, Peck KR, Chongthaleong A, Lalitha MK, Perera J, Yee TT, JAMAl F, Kamarulzaman A, Carlos CC, So T. 2004. Clinical outcomes of pneumococcal pneumonia caused by antibiotic-resistant strains in asian countries: a study by the Asian Network for Surveillance of Resistant Pathogens. Clin Infect Dis 38(11):1570-8.

112. Ekdahl K, Martensson A, Kamme C. 1998. Bacteraemic pneumococcal infections in Southern Sweden 1981-96: trends in incidence, mortality, age-distribution, serogroups and penicillin-resistance. Scand J Infect Dis 30(3):257-62.

113. Altclas J, Salgueira C, Di Martino A. 2004. Pneumococcal bacteremia in a single center in Argentina. Int J Infect Dis 8(1):53-8.

114. Aspa J, Rajas O, Rodriguez de Castro F, Huertas MC, Borderias L, Cabello FJ, Tabara J, Hernandez-Flix S, Martinez-Sanchis A, Torres A. 2006. Impact of initial antibiotic choice on mortality from pneumococcal pneumonia. Eur Respir J 27(5):1010-9.

115. Sangthawan P, Chantaratchada S, Chanthadisai N, Wattanathum A. 2003. Prevalence and clinical significance of community-acquired penicillin-resistant pneumococcal pneumonia in Thailand. Respirology 8(2):208-12.

116. Aspa J, Rajas O, Rodriguez de Castro F, Blanquer J, Zalacain R, Fenoll A, de Celis R, Vargas A, Rodriguez Salvanes F, Espana PP, Rello J, Torres A. 2004. Drug-

resistant pneumococcal pneumonia: clinical relevance and related factors. Clin Infect Dis 38(6):787-98.

117. Fine MJ, Pratt HM, Obrosky DS, Lave JR, McIntosh LJ, Singer DE, Coley CM, Kapoor WN. 2000. Relation between length of hospital stay and costs of care for patients with community-acquired pneumonia. Am J Med 109(5):378-85.

118. Klepser ME, Klepser DG, Ernst EJ, Brooks J, Diekema DJ, Mozaffari E, Hendrickson J, Doern GV. 2003. Health care resource utilization associated with treatment of penicillin-susceptible and -nonsusceptible isolates of Streptococcus pneumoniae. Pharmacotherapy 23(3):349-59.

119. Niederman MS, McCombs JS, Unger AN, Kumar A, Popovian R. 1998. The cost of treating community-acquired pneumonia. Clin Ther 20(4):820-37.

120. Pallares R, Linares J, Vadillo M, Cabellos C, Manresa F, Viladrich PF, Martin R, Gudiol F. 1995. Resistance to penicillin and cephalosporin and mortality from severe pneumococcal pneumonia in Barcelona, Spain. N Engl J Med 333(8):474-80.

121. Birnbaum HG, Morley M, Greenberg PE, Colice GL. 2002. Economic burden of respiratory infections in an employed population. Chest 122(2):603-11.

122. Guest JF, Morris A. 1997. Community-acquired pneumonia: the annual cost to the National Health Service in the UK. Eur Respir J 10(7):1530-4.

123. Tan TQ, Mason EO, Jr., Barson WJ, Wald ER, Schutze GE, Bradley JS, Arditi M, Givner LB, Yogev R, Kim KS, Kaplan SL. 1998. Clinical characteristics and outcome of children with pneumonia attributable to penicillin-susceptible and penicillin-nonsusceptible Streptococcus pneumoniae. Pediatrics 102(6):1369-75.

124. Shah PB, Giudice JC, Griesback R, Jr., Morley TF, Vasoya A. 2004. The newer guidelines for the management of community-acquired pneumonia. J Am Osteopath Assoc 104(12):521-6.

125. Kaplan V, Angus DC, Griffin MF, Clermont G, Scott Watson R, Linde-Zwirble WT. 2002. Hospitalized community-acquired pneumonia in the elderly: age- and sex-related patterns of care and outcome in the United States. Am J Respir Crit Care Med 165(6):766-72.

126. Centers for Disease Control and Prevention. 1997. Prevention of pneumococcal disease: recommendations of the Advisory Committee on Immunization Practices (ACIP). MMWR 46(No. RR-8).

127. Centers for Disease Control and Prevention. National Vital Statistics System. 2011 Oct 19]. Available from: http://205.207.175.93/VitalStats/ReportFolders/reportFolders.aspx

128. U.S. Census Bureau. Annual Estimates of the Resident Population by Sex and Selected Age Groups for the United States: April 1, 2000 to July 1, 2009 (NC-EST2009-02). 2011 Oct 19]. Available from: http://www.census.gov/popest/national/asrh/NC-EST2009-sa.html

129. Centers for Disease Control and Prevention. 2010. Recommended immunization schedules for persons aged 0 through 18 years--United States. MMWR Morb Mortal Wkly Rep 58(51&52).

130. Centers for Disease Control and Prevention. Datasets and Related Documentation for the National Immunization Survey, 2005 to Present. 2011 Oct 19]. Available from: http://www.cdc.gov/nchs/nis/data_files.htm

131. Foster JE. 2008. Nosocomial Pneumonia. In: Myers JA, Millikan KW, Sacalarides TJ, editors. Common surgical diseases : an algorithmic approach to problem solving. Second ed. New York: Springer.

132. Brandenburg JA, Marrie TJ, Coley CM, Singer DE, Obrosky DS, Kapoor WN, Fine MJ. 2000. Clinical presentation, processes and outcomes of care for patients with pneumococcal pneumonia. J Gen Intern Med 15(9):638-46.

133. Mandell L. 2005. Decisions about treating community-acquired pneumonia. Ann Intern Med 142(3):215-6.

134. Feikin DR, Schuchat A, Kolczak M, Barrett NL, Harrison LH, Lefkowitz L, McGeer A, Farley MM, Vugia DJ, Lexau C, Stefonek KR, Patterson JE, Jorgensen JH. 2000. Mortality from invasive pneumococcal pneumonia in the era of antibiotic resistance, 1995-1997. Am J Public Health 90(2):223-9.

135. Arditi M, Mason EO, Jr., Bradley JS, Tan TQ, Barson WJ, Schutze GE, Wald ER, Givner LB, Kim KS, Yogev R, Kaplan SL. 1998. Three-year multicenter surveillance of pneumococcal meningitis in children: clinical characteristics, and outcome related to penicillin susceptibility and dexamethasone use. Pediatrics 102(5):1087-97.

136. Huang SS, Johnson KM, Ray GT, Wroe P, Lieu TA, Moore MR, Zell ER, Linder JA, Grijalva CG, Metlay JP, Finkelstein JA. 2011. Healthcare utilization and cost of pneumococcal disease in the United States. Vaccine 29(18):3398-412.

137. Centers for Disease Control and Prevention. Catch-up Immunization Schedule for Persons Aged 4 Months Through 18 Years Who Start Late or Who Are More Than 1 Month Behind. 2011 Oct 19]. Available from: http://www.cdc.gov/vaccines/recs/schedules/downloads/child/catchup-schedule-pr.pdf

138. Guevara RE, Butler JC, Marston BJ, Plouffe JF, File TM, Jr., Breiman RF. 1999. Accuracy of ICD-9-CM codes in detecting community-acquired pneumococcal pneumonia for incidence and vaccine efficacy studies. Am J Epidemiol 149(3):282-9.

APPENDICES

Appendix A
Equations for the Dynamic Transmission Model

The differential equations used in the model are shown below. The *i* subscript indicates age group (1, 2, or 3). Subscripts for IPD also include 1 for VT and 2 for NVT, before the age subscript. Parameters g_0, g_3, and v_3 were included for brevity of presentation, and were set equal to zero. $\delta_{ij} = 1$ when $i = j$ and 0 when $i \neq j$.

$$dN_{SSi}/dt = \delta_{i1}\mu_N N_3 + g_{i-1}N_{SS(i-1)} + \delta_{i3}g_{i-1}N_{VSS(i-1)} + \rho_1 N_{RSi} + \rho_2 N_{SRi} - (\lambda_{1i} + \lambda_{2i})N_{SSi} - (v_i + g_i + \mu_i)N_{SSi}$$
$$dN_{SIi}/dt = g_{i-1}N_{SI(i-1)} + \delta_{i3}g_{i-1}N_{VSI(i-1)} + \lambda_{2i}N_{SSi} + \rho_1 N_{RIi} - (c_1\lambda_{1i} + \gamma_{2i})N_{SIi} - (v_i + g_i + \mu_i)N_{SIi}$$
$$dN_{SRi}/dt = g_{i-1}N_{SR(i-1)} + \delta_{i3}g_{i-1}N_{VSR(i-1)} + \gamma_{2i}N_{SIi} + \rho_1 N_{RRi} - (\lambda_{1i} + \rho_2)N_{SRi} - (v_i + g_i + \mu_i)N_{SRi}$$
$$dN_{ISi}/dt = g_{i-1}N_{IS(i-1)} + \delta_{i3}g_{i-1}N_{VIS(i-1)} + \lambda_{1i}N_{SSi} + \rho_2 N_{IRi} - (c_2\lambda_{2i} + \gamma_{1i})N_{ISi} - (v_i + g_i + \mu_i)N_{ISi}$$
$$dN_{IIi}/dt = g_{i-1}N_{II(i-1)} + \delta_{i3}g_{i-1}N_{VII(i-1)} + c_1\lambda_{1i}N_{SIi} + c_2\lambda_{2i}N_{ISi} - (\gamma_{1i} + \gamma_{2i})N_{IIi} - (v_i + g_i + \mu_i)N_{IIi}$$
$$dN_{IRi}/dt = g_{i-1}N_{IR(i-1)} + \delta_{i3}g_{i-1}N_{VIR(i-1)} + \lambda_{1i}N_{SRi} + \gamma_{2i}N_{IIi} - (\rho_2 + \gamma_{1i})N_{IRi} - (v_i + g_i + \mu_i)N_{IRi}$$
$$dN_{RSi}/dt = g_{i-1}N_{RS(i-1)} + \delta_{i3}g_{i-1}N_{VRS(i-1)} + \gamma_{1i}N_{ISi} + \rho_2 N_{RRi} - (\rho_1 + \lambda_{2i})N_{RSi} - (v_i + g_i + \mu_i)N_{RSi}$$
$$dN_{RIi}/dt = g_{i-1}N_{RI(i-1)} + \delta_{i3}g_{i-1}N_{VRI(i-1)} + \lambda_{2i}N_{RSi} + \gamma_{1i}N_{IIi} - (\rho_1 + \gamma_{2i})N_{RIi} - (v_i + g_i + \mu_i)N_{RIi}$$
$$dN_{RRi}/dt = g_{i-1}N_{RR(i-1)} + \delta_{i3}g_{i-1}N_{VRR(i-1)} + \gamma_{1i}N_{IRi} + \gamma_{2i}N_{RIi} - (\rho_1 + \rho_2)N_{RRi} - (v_i + g_i + \mu_i)N_{RRi}$$
$$dN_{VSSi}/dt = (1-\delta_{i3})g_{i-1}N_{VSS(i-1)} + v_i N_{SSi} + \rho_1 N_{VRSi} + \rho_2 N_{VSRi} - (e\lambda_{1i} + k\lambda_{2i})N_{VSSi} - (g_i + \mu_i)N_{VSSi}$$
$$dN_{VSIi}/dt = (1-\delta_{i3})g_{i-1}N_{VSI(i-1)} + v_i N_{SIi} + k\lambda_{2i}N_{VSSi} + \rho_1 N_{VRIi} - (ec_1\lambda_{1i} + \gamma_{2i})N_{VSIi} - (g_i + \mu_i)N_{VSIi}$$
$$dN_{VSRi}/dt = (1-\delta_{i3})g_{i-1}N_{VSR(i-1)} + v_i N_{SRi} + \gamma_{2i}N_{VSIi} + \rho_1 N_{VRRi} - (e\lambda_{1i} + \rho_2)N_{VSRi} - (g_i + \mu_i)N_{VSRi}$$
$$dN_{VISi}/dt = (1-\delta_{i3})g_{i-1}N_{VIS(i-1)} + v_i N_{ISi} + e\lambda_{1i}N_{VSSi} + \rho_2 N_{VIRi} - (kc_2\lambda_{2i} + \gamma_{1i})N_{VISi} - (g_i + \mu_i)N_{VISi}$$
$$dN_{VIIi}/dt = (1-\delta_{i3})g_{i-1}N_{VII(i-1)} + v_i N_{IIi} + ec_1\lambda_{1i}N_{VSIi} + kc_2\lambda_{2i}N_{VISi} - (\gamma_{1i} + \gamma_{2i})N_{VIIi} - (g_i + \mu_i)N_{VIIi}$$
$$dN_{VIRi}/dt = (1-\delta_{i3})g_{i-1}N_{VIR(i-1)} + v_i N_{IRi} + e\lambda_{1i}N_{VSRi} + \gamma_{2i}N_{VIIi} - (\rho_2 + \gamma_{1i})N_{VIRi} - (g_i + \mu_i)N_{VIRi}$$
$$dN_{VRSi}/dt = (1-\delta_{i3})g_{i-1}N_{VRS(i-1)} + v_i N_{RSi} + \gamma_{1i}N_{VISi} + \rho_2 N_{VRRi} - (\rho_1 + k\lambda_{2i})N_{VRSi} - (g_i + \mu_i)N_{VRSi}$$
$$dN_{VRIi}/dt = (1-\delta_{i3})g_{i-1}N_{VRI(i-1)} + v_i N_{RIi} + k\lambda_{2i}N_{VRSi} + \gamma_{1i}N_{VIIi} - (\rho_1 + \gamma_{2i})N_{VRIi} - (g_i + \mu_i)N_{VRIi}$$
$$dN_{VRRi}/dt = (1-\delta_{i3})g_{i-1}N_{VRR(i-1)} + v_i N_{RRi} + \gamma_{1i}N_{VIRi} + \gamma_{2i}N_{VRIi} - (\rho_1 + \rho_2)N_{VRRi} - (g_i + \mu_i)N_{VRRi}$$
$$dN_{IPD1i}/dt = f_{1i}\lambda_{1i}(N_{SSi} + c_1 N_{SIi} + N_{SRi} + er(N_{VSSi} + c_1 N_{VSIi} + N_{VSRi}))$$
$$dN_{IPD2i}/dt = f_{2i}\lambda_{2i}(N_{SSi} + c_2 N_{ISi} + N_{RSi} + k(N_{VSSi} + c_2 N_{VISi} + N_{VRSi}))$$

where:

$$\lambda_{1i} = \sum_{j=1}^{3} \frac{\beta_{1ij}(N_{ISj} + N_{IIj} + N_{IRj} + N_{VISj} + N_{VIIj} + N_{VIRj})}{N_j}$$
$$\lambda_{2i} = \sum_{j=1}^{3} \frac{\beta_{1ij}(N_{SIj} + N_{IIj} + N_{RIj} + N_{VSIj} + N_{VIIj} + N_{VRIj})}{N_j}$$

Appendix B
U. S. Population Estimates and Predictions

Table B.1
Annual Population Estimates

Year	<2 years old	2-15 years old	≥16 years old	Live Births	Number of Infant Deaths
1999	7,576,800	54,802,500	210,433,000	4,158,212	27,937
2000	8,305,243	56,643,801	217,810,639	4,058,814	28,035
2001	8,458,885	56,776,038	220,446,741	4,025,933	27,568
2002	8,546,362	56,923,853	222,968,177	4,021,726	28,034
2003	8,617,372	56,984,993	225,281,695	4,089,950	28,025
2004	8,752,964	56,980,966	227,821,676	4,112,052	27,936
2005	8,849,581	56,929,430	230,411,492	4,138,349	28,440
2006	8,902,216	57,025,537	233,246,527	4,265,555	28,527
2007	9,057,285	57,147,620	236,016,362	4,316,233	29,138
2008	9,169,199	57,274,704	238,597,183	4,247,694	28,029
2009	8,519,862	57,659,087	240,989,961	4,131,019	26,526

Table B.2
Projected Population Size and Vaccination Through 2025

Year	Predicted Population Size, <2 years old	Predicted Population Size, 2-15 years old	Predicted Population Size, ≥16 years old	Predicted Number of Births	Predicted Number of Infant Deaths	Predicted Number Eligible for Vacc.	Predicted Number of Doses Administered (1000s)	Predicted Vaccine Cost[a]
2010	9,209,265	57,817,272	244,456,296	4,239,271	28,500	4,210,800	15,290	$1,745
2011	9,308,450	57,981,536	247,259,263	4,251,832	28,600	4,223,200	15,330	$1,750
2012	9,407,634	58,145,801	250,062,230	4,263,576	28,700	4,234,900	15,370	$1,754
2013	9,506,818	58,310,065	252,865,196	4,274,502	28,800	4,245,700	15,410	$1,759
2014	9,606,002	58,474,330	255,668,163	4,284,611	28,900	4,255,700	15,450	$1,764
2015	9,705,186	58,638,594	258,471,130	4,293,903	29,000	4,264,900	15,480	$1,767
2016	9,804,370	58,802,859	261,274,097	4,302,377	29,100	4,273,300	15,510	$1,770
2017	9,903,554	58,967,123	264,077,063	4,310,035	29,200	4,280,800	15,540	$1,774
2018	10,002,738	59,131,388	266,880,030	4,316,874	29,300	4,287,600	15,560	$1,776
2019	10,101,922	59,295,652	269,682,997	4,322,897	29,300	4,293,600	15,590	$1,780
2020	10,201,106	59,459,917	272,485,964	4,328,102	29,400	4,298,700	15,610	$1,782
2021	10,300,291	59,624,181	275,288,930	4,332,491	29,500	4,303,000	15,620	$1,783
2022	10,399,475	59,788,446	278,091,897	4,336,061	29,500	4,306,600	15,630	$1,784
2023	10,498,659	59,952,710	280,894,864	4,338,815	29,600	4,309,200	15,640	$1,785
2024	10,597,843	60,116,975	283,697,830	4,340,751	29,600	4,311,200	15,650	$1,786
2025	10,697,027	60,281,240	286,500,797	4,341,870	29,700	4,312,200	15,650	$1,786
Total	--	--	--	--	--	--	248,330	$28,345

a Costs are given in millions of dollars.

Appendix C Estimates from the Dynamic Transmission Model

Table C.1

Model-based estimates of fractions of subjects vaccinated and colonized, by age group, serotype group, and year

	Vaccinated [a]		VT Colonization, Vacc. Subjects [a]		VT Colonization, Unvaccinated Subjects			VT Colonization, All Subjects			NVT Colonization, All Subjects			Total Colonization [b], All Subjects		
Year	<2	2-15	<2	2-15	<2	2-15	≥16	<2	2-15	≥16	<2	2-15	≥16	<2	2-15	≥16
1999	0	0	0	0	0.2954	0.1459	0.0431	0.2954	0.1459	0.0431	0.3072	0.1478	0.0451	0.5179	0.2751	0.0867
2000	0.1008	0.0023	0.0256	0.0003	0.263	0.1449	0.0429	0.2886	0.1452	0.0429	0.3074	0.1478	0.0451	0.5132	0.2746	0.0864
2001	0.6815	0.0342	0.1045	0.0029	0.0601	0.1099	0.031	0.1645	0.1128	0.031	0.3164	0.1509	0.0461	0.4331	0.2492	0.0759
2002	0.8581	0.1012	0.0575	0.0038	0.0091	0.0569	0.0153	0.0666	0.0606	0.0153	0.3271	0.1565	0.0478	0.3739	0.2091	0.0625
2003	0.8987	0.1669	0.0255	0.0026	0.0023	0.0252	0.0069	0.0278	0.0278	0.0069	0.3316	0.1596	0.0487	0.3511	0.1837	0.0553
2004	0.9158	0.2187	0.011	0.0014	0.0007	0.0105	0.003	0.0117	0.012	0.003	0.3334	0.161	0.049	0.3416	0.1714	0.0519
2005	0.9301	0.2676	0.0047	0.0007	0.0002	0.0041	0.0012	0.0049	0.0049	0.0012	0.3341	0.1617	0.0492	0.3376	0.1659	0.0504
2006	0.9426	0.3142	0.0019	0.0003	0.0001	0.0016	0.0005	0.002	0.0019	0.0005	0.3345	0.1619	0.0493	0.3358	0.1636	0.0497
2007	0.9522	0.3582	0.0008	0.0001	0	0.0006	0.0002	0.0008	0.0007	0.0002	0.3346	0.162	0.0493	0.3351	0.1626	0.0495
2008	0.9573	0.3997	0.0003	0.0001	0	0.0002	0.0001	0.0003	0.0003	0.0001	0.3347	0.1621	0.0493	0.3349	0.1623	0.0494
2009	0.9573	0.4385	0.0001	0	0	0.0001	0	0.0001	0.0001	0	0.3347	0.1621	0.0493	0.3347	0.1622	0.0493
2010	0.9573	0.4745	0	0	0	0	0	0	0	0	0.3347	0.1621	0.0493	0.3347	0.1621	0.0493
2011	0.9573	0.5081	0	0	0	0	0	0	0	0	0.3347	0.1621	0.0493	0.3347	0.1621	0.0493
2012	0.9573	0.5393	0	0	0	0	0	0	0	0	0.3347	0.1621	0.0493	0.3347	0.1621	0.0493
2013	0.9573	0.5684	0	0	0	0	0	0	0	0	0.3347	0.1621	0.0493	0.3347	0.1621	0.0493
2014	0.9573	0.5954	0	0	0	0	0	0	0	0	0.3347	0.1621	0.0493	0.3347	0.1621	0.0493
2015	0.9573	0.6206	0	0	0	0	0	0	0	0	0.3347	0.1621	0.0493	0.3347	0.1621	0.0493
2016	0.9573	0.644	0	0	0	0	0	0	0	0	0.3347	0.1621	0.0493	0.3347	0.1621	0.0493
2017	0.9573	0.6658	0	0	0	0	0	0	0	0	0.3347	0.1621	0.0493	0.3347	0.1621	0.0493
2018	0.9573	0.686	0	0	0	0	0	0	0	0	0.3347	0.1621	0.0493	0.3347	0.1621	0.0493
2019	0.9573	0.7049	0	0	0	0	0	0	0	0	0.3347	0.1621	0.0493	0.3347	0.1621	0.0493
2020	0.9573	0.7224	0	0	0	0	0	0	0	0	0.3347	0.1621	0.0493	0.3347	0.1621	0.0493
2021	0.9573	0.7388	0	0	0	0	0	0	0	0	0.3347	0.1621	0.0493	0.3347	0.1621	0.0493
2022	0.9573	0.754	0	0	0	0	0	0	0	0	0.3347	0.1621	0.0493	0.3347	0.1621	0.0493
2023	0.9573	0.7681	0	0	0	0	0	0	0	0	0.3347	0.1621	0.0493	0.3347	0.1621	0.0493

Table C.1
Continued

	Vaccinated [a]		VT Colonization, Vacc. Subjects [a]		VT Colonization, Unvaccinated Subjects			VT Colonization, All Subjects			NVT Colonization, All Subjects			Total Colonization [b], All Subjects		
Year	<2	2-15	<2	2-15	<2	2-15	≥16	<2	2-15	≥16	<2	2-15	≥16	<2	2-15	≥16
2024	0.9573	0.7813	0	0	0	0	0	0	0	0	0.3347	0.1621	0.0493	0.3347	0.1621	0.0493
2025	0.9573	0.7935	0	0	0	0	0	0	0	0	0.3347	0.1621	0.0493	0.3347	0.1621	0.0493

a Vaccination effects wane by adulthood.
b Total fraction of subjects who were colonized with VT, NVT, or both.

Table C.2
Model-based estimates of reported IPD incidence per 100,000 people, by age group, serotype group, and year

	VT IPD			NVT IPD			Total Reported IPD			
Year	<2	2-15	≥16	<2	2-15	≥16	<2	2-15	≥16	All ages
1999	188.5	5.9	10.7	3.8	5.2	16.9	192.3	11	27.6	28.7
2000	158.8	5.7	10.3	3.8	5.2	16.9	162.5	10.9	27.3	27.7
2001	50.1	4	7.1	3.9	5.3	17.4	53.9	9.4	24.5	22.8
2002	12.2	2	3.5	4	5.5	18	16.2	7.5	21.4	19.3
2003	4.2	0.9	1.5	4.1	5.6	18.3	8.2	6.5	19.8	17.6
2004	1.6	0.4	0.7	4.1	5.6	18.4	5.7	6	19.1	16.9
2005	0.6	0.1	0.3	4.1	5.7	18.5	4.7	5.8	18.7	16.5
2006	0.2	0.1	0.1	4.1	5.7	18.5	4.3	5.7	18.6	16.4
2007	0.1	0	0	4.1	5.7	18.5	4.2	5.7	18.5	16.3
2008	0	0	0	4.1	5.7	18.5	4.1	5.7	18.5	16.3
2009	0	0	0	4.1	5.7	18.5	4.1	5.7	18.5	16.3
2010	0	0	0	4.1	5.7	18.5	4.1	5.7	18.5	16.3
2011	0	0	0	4.1	5.7	18.5	4.1	5.7	18.5	16.3
2012	0	0	0	4.1	5.7	18.5	4.1	5.7	18.5	16.3
2013	0	0	0	4.1	5.7	18.5	4.1	5.7	18.5	16.3
2014	0	0	0	4.1	5.7	18.5	4.1	5.7	18.5	16.3
2015	0	0	0	4.1	5.7	18.5	4.1	5.7	18.5	16.3
2016	0	0	0	4.1	5.7	18.5	4.1	5.7	18.5	16.3
2017	0	0	0	4.1	5.7	18.5	4.1	5.7	18.5	16.3
2018	0	0	0	4.1	5.7	18.5	4.1	5.7	18.5	16.3
2019	0	0	0	4.1	5.7	18.5	4.1	5.7	18.5	16.3
2020	0	0	0	4.1	5.7	18.5	4.1	5.7	18.5	16.3
2021	0	0	0	4.1	5.7	18.5	4.1	5.7	18.5	16.3
2022	0	0	0	4.1	5.7	18.5	4.1	5.7	18.5	16.3

Table C.2
Continued

	VT IPD			NVT IPD			Total Reported IPD			
Year	<2	2-15	≥16	<2	2-15	≥16	<2	2-15	≥16	All ages
2023	0	0	0	4.1	5.7	18.5	4.1	5.7	18.5	16.3
2024	0	0	0	4.1	5.7	18.5	4.1	5.7	18.5	16.3
2025	0	0	0	4.1	5.7	18.5	4.1	5.7	18.5	16.3

Note: Estimates of IPD were developed using reported IPD figures from CDC ABCs, which primarily include meningitis and bacteremia.