



## Practice of Epidemiology

# Relating Pneumococcal Carriage Among Children to Disease Rates Among Adults Before and After the Introduction of Conjugate Vaccines

Daniel M. Weinberger\*, Lindsay R. Grant, Robert C. Weatherholtz, Joshua L. Warren, Katherine L. O'Brien, and Laura L. Hammitt

\* Correspondence to Dr. Daniel M. Weinberger, Department of Epidemiology of Microbial Diseases, Yale School of Public Health, P.O. Box 208034, New Haven, CT 06520 (e-mail: Daniel.weinberger@yale.edu).

Initially submitted May 1, 2015; accepted for publication October 13, 2015.

The use of pneumococcal conjugate vaccines (PCVs) in children has a strong indirect effect on disease rates in adults. When children are vaccinated with PCVs, other serotypes that are not targeted by the vaccine can increase in frequency (serotype replacement) and reduce the direct and indirect benefits of the vaccine. To understand and predict the likely impacts of serotype replacement, it is important to know how patterns in the transmission of serotypes among children relate to disease rates in adults. We used data on pneumococcal carriage and disease from Navajo Nation children and adults collected before and after the routine use of PCVs (1998–2012). Using regression models within a Bayesian framework, we found that serotype-specific carriage and invasiveness (disease incidence divided by carriage prevalence) had similar patterns in children and adults. Moreover, carriage in children, invasiveness in children, and a serotype-specific random intercept (which captured additional variation associated with the serotypes) could predict the incidence serotype-specific pneumococcal disease in adults 18–39 years of age and those 40 years of age or older in the era of routine use of PCVs. These models could help us predict the effects of future pneumococcal vaccine use in children on disease rates in adults, and the modeling approach developed here could be used to test these findings in other settings.

adults; carriage; colonization; comorbid conditions; conjugate vaccines; pneumococcus; *Streptococcus pneumoniae*

Abbreviations: CI, confidence interval; CrI, credible interval; IPD, invasive pneumococcal disease; PCV, pneumococcal conjugate vaccine.

Pneumococcal conjugate vaccines (PCVs) have been used for 15 years in children and were recently recommended for use in adults 65 years of age or older in the United States (1). Immunizing children with the first licensed PCV (PCV7) prevented disease caused by the targeted serotypes and reduced nasopharyngeal transmission of these serotypes (2, 3). Because children are a major source of pneumococcal transmission to persons of all ages (4, 5), vaccinating children also led to the near-elimination of vaccine-targeted serotypes as causes of invasive pneumococcal disease (IPD) among adults (3). However, the net impact of PCVs is the sum of both the decline in disease caused by PCV-targeted serotypes and the increase in the incidence of IPD caused by nontargeted serotypes (i.e., serotype replacement) (2, 3, 6, 7). Although the increase in the rate of IPD caused by serotypes not targeted

by the PCV was small relative to the declines in IPD caused by serotypes that were targeted in children, this increase substantially offset the declines in PCV-targeted IPD in some adult age groups and settings (3).

Expanded valency PCVs are now available (PCV10 and PCV13), and PCVs that target additional serotypes (e.g., 15 or more serotypes) are under development; however, there are at least 80 serotypes not targeted by these vaccines, and little is known about how increases in carriage prevalence of different serotypes would impact disease rates in adults. Forecasting how the use of these expanded-serotype PCVs might impact disease rates in older adults depends on clearly understanding the relationship between carriage and disease (7).

In children, the rate of disease caused by a serotype can be estimated by combining the serotype-specific carriage

frequency and the “invasiveness” of the serotype, which is defined as the rate of disease per nasopharyngeal acquisition and often approximated by the colonization point prevalence of a serotype (2, 7–10). Among children, the invasiveness of serotypes is conserved over time and across diverse geographic settings (11). Differences in invasiveness are biologically determined by the capsule and other bacterial factors (12–14) and can be linked with the ability of the serotypes to avoid host immunity (13).

In adults, the link between carriage and disease rates has not been clearly defined. In a study from the United Kingdom, (15) Trotter et al. found broadly similar invasiveness values by serotype between those  $\geq 5$  and  $< 5$  years of age. However, detailed breakdowns by age and underlying conditions were not available. We hypothesized that invasiveness values would be similar in children and in adults without risk factors for pneumococcal disease. Moreover, we hypothesized that in adults with risk factors for pneumococcal disease, carriage prevalence (exposure) would determine serotype patterns of IPD but differences in invasiveness between serotypes would be less important because host susceptibility would outweigh differences in bacterial virulence.

In the present analysis, we used carriage and disease data from Navajo and White Mountain Apache children and adults, populations with high rates of pneumococcal disease (16–18). We compared the carriage patterns in children and adults, estimated and compared the invasiveness of the most common serotypes

among children and adults, and then fit and tested regression models to link the frequency of serotypes being carried by children with the frequency of serotypes causing disease in adults with and without risk factors for pneumococcal disease.

## METHODS

### Data sources

Nasopharyngeal carriage and invasive disease data were obtained from studies conducted among the Navajo and White Mountain Apache populations in the southwestern United States (Table 1). The details of these study objectives and designs have been previously described (18–20) and are summarized here briefly. There were 4 community-based carriage studies focused on the evaluation of PCV effects among children. PCV7 was introduced as a routine pediatric vaccine in October 2000 and was replaced by PCV13 in April 2010. A catch-up campaign was conducted among children younger than 5 years of age in both instances. Study 1 (1998–2000) was a randomized controlled trial designed to evaluate PCV7 efficacy against invasive pneumococcal disease, as well as the direct and indirect effects of PCV7 on nasopharyngeal colonization in vaccine recipients and their siblings younger than 6 years of age (20). We used only the carriage data from the controls. In study 2 (2001–2002), investigators examined nasopharyngeal carriage in children and adults in those same

**Table 1.** Carriage and Invasive Pneumococcal Disease Study Characteristics, Navajo and White Mountain Apache Populations, United States, 1995–2012

Age, years	Period <sup>a</sup>	Dates <sup>b</sup>	Colonization				No. of IPD Cases	Reference No. <sup>c</sup>
			No. of Participants	Total No. of Swabs	No. of Positive Swabs	%		
<b>&lt;7</b>								
	1	June 1995–June 2000	411	1,077	684	63.5	252	20
	2	July 2000–June 2004	1,914	1,918	1,185	61.8	117	19
	3	July 2004–June 2010	549	3,329	1,835	55.1	132	18
	4	July 2010–December 2012	3,259	3,259	1,686	51.7	22	
<b>18–39</b>								
	1	June 1995–June 2000	N/A	N/A	N/A	N/A	116	N/A
	2	July 2000–June 2004	1,494	1,507	180	11.9	56	19
	3	July 2004–June 2010	373	2,199	240	10.9	96	18
	4	July 2010–December 2012	2,307	2,307	277	12.0	24	N/A
<b><math>\geq 40</math></b>								
	1	June 1995–June 2000	N/A	N/A	N/A	N/A	311	N/A
	2	July 2000–June 2004	500	504	48	9.5	205	19
	3	July 2004–June 2010	67	427	41	9.6	418	18
	4	July 2010–December 2012	399	399	37	9.3	110	N/A

Abbreviations: IPD, invasive pneumococcal disease; N/A, not applicable.

<sup>a</sup> Carriage data were collected as part of a randomized controlled trial, and only data from the control group (who did not receive pneumococcal conjugate vaccine serotype 7) were included in this analysis.

<sup>b</sup> Date ranges denote when the disease isolates were collected. Carriage isolates were collected over shorter periods of time within these windows (1998–2000, 2001–2002, 2006–2008, and 2010–2012).

<sup>c</sup> Carriage data from period 4 are unpublished.

communities shortly after the licensure of PCV7 (19). Study 3 (2006–2008) was designed to examine intrahousehold transmission among children and adults in these communities, with 7 nasopharyngeal samples collected from each person over a 6-month period several years after the introduction of PCV7 (18). In study 4 (2010–2012), researchers examined pneumococcal carriage among children and adults after the introduction of PCV13 (K.L.O., unpublished data, 2015). In the analyses presented here, we focused on children younger than 7 years of age, adults 18–39 years of age, and adults 40 years of age or older. These age groups were chosen based on age patterns in the carriage data. The carriage prevalence among Navajo children is stable among children between the ages of 1 and 7 years but then declines rapidly. Among adults, the majority of the samples were collected among those aged 18–39 years, and this represents the ages at which people are most likely to be parents of young children. The remaining samples in adults included individuals ranging in age from their 40s to their 90s, but the numbers did not allow for further stratification.

The number of cases of IPD that occurred during each time period was derived from population-based active bacterial surveillance conducted at health facilities serving the Navajo and White Mountain Apache reservations, as described previously (17) (1995–2012). For analyses stratified by comorbid conditions, cases were divided into 2 groups based on whether or not the individual had at least 1 previously recognized comorbid condition that is included on the Charlson comorbidity index (21). The comorbid conditions included congestive heart failure, atherosclerosis, chronic obstructive pulmonary disease, diabetes, renal disease, leukemia, lymphoma, metastatic cancer, tumor without metastasis, cirrhosis, and acquired immunodeficiency syndrome. We were missing serotype information for some cases of IPD; the degree of missingness differing by time period (19% for the period before PCV7 and 10%–12% for the period after PCV7 and PCV13). For case patients for whom serotype information was missing, we assumed that the serotype distribution was the same as that in case patients with complete serotype information in the same time period and age group (and comorbidity group). This assumption is reasonable because missingness typically resulted from sample handling issues and was therefore unlikely to be biased. For all studies, the serotypes of the bacterial isolates were determined based on serological reactions. Disease incidence was defined as the number of cases per 1,000 people per year.

All studies received ethical approval from the Institutional Review Board at Johns Hopkins Bloomberg School of Public Health, the Navajo Nation Institutional Review Board, the Indian Health Service Phoenix Area Institutional Review Board, and the Navajo and White Mountain Apache tribes. Parents or guardians of children enrolled in the carriage studies provided written informed consent, as did adults who participated in the carriage studies.

### Estimating uncertainty in serotype-specific carriage prevalence and invasiveness

Estimates of serotype-specific carriage prevalence and invasiveness were based on sparse data. We estimated both quantities in a Bayesian framework to obtain robust estimates of the values and their uncertainty. The details of these

procedures are described Web Appendix 1 (available at <http://aje.oxfordjournals.org/>).

### Predicting disease rates in different age groups as a function of carriage and invasiveness in children

We sought to develop a model that directly links carriage prevalence in children (for whom the data are more robust) with disease incidence in adults. Such a model would also be epidemiologically relevant because children are thought to be a major source of pneumococcal exposure for adults. In order to validate such a model linking pediatric carriage and adult disease, we fit the models to adult IPD data from 3 of the 4 available time periods and then inserted the observed pediatric carriage prevalence from the fourth time period into the model to predict the adult IPD incidence in the fourth time period. We then repeated this process iteratively, fitting to a different set of 3 time periods and predicting IPD incidence in the fourth time period. The models were fit as Poisson regressions (separately by age group) in which the outcome variable was the incidence of IPD caused by each serotype in each period for a given age group. The covariates were  $\log(\text{carriage prevalence})$  among children,  $\log(\text{invasiveness})$  among children, a random effect for serotype, and a random effect for individual measurements to capture overdispersion (22). For serotype  $i$ , and time period  $k$ ,

$$\begin{aligned} \text{Disease rate}_{ik} = & \exp(b_0 + b_1 * \log\_true\_carriage_{ik} \\ & + b_2 * \log\_true\_invasiveness_i + \text{serotype}_i \\ & + \text{observation}_{ik}). \end{aligned}$$

To account for measurement error in carriage prevalence and invasiveness, we modeled the true, unobserved values in a Bayesian framework, as described in Web Appendix 1, and included these modeled estimates as covariates in the regression. The posterior estimates of invasiveness (mean and precision) from these previous models were used as priors for “ $\log\_true\_invasiveness_i$ ” and “ $\log\_true\_carriage$ .” All of the priors were chosen to be weakly informative; for the regression coefficients, normal priors were used (mean = 0, precision =  $1 \times 10^{-5}$ ). The serotype-specific random effect and the observation-level random effects had independent normal priors (mean = 0; precision had independent  $\gamma(0.001, 0.001)$  priors). We also considered simpler models that did not include the random effects for serotype or observation but used the full model because of better deviance information criterion scores or improved predictions in the cross-validation analyses (see Web Appendix 1 for details).

Models were fit using the RJags package (23) in R (24). The code used for the cross-validations and sample data (with some random noise added to the true values for privacy reasons) are included in Web Appendix 2. Additionally, we have included code to generate predictions of serotype-specific IPD cases in adults using some sample carriage data (Web Code).

## RESULTS

### Overview of data

There were 523 cases of IPD among children younger than 7 years of age, 292 cases in adults 18–39 years of age, and 1,044

cases in adults 40 years of age or older in 1995–2012 (Table 1). Among adults 40 years of age or older, the median age of case patients was 60 years (range, 40–96 years). During the 4 carriage surveys, the average pneumococcal carriage prevalence was 5,390 out of 9,583 (56.2%) among children younger than 7 years of age, 697 out of 6,013 (11.6%) among adults 18–39 years of age, and 89 out of 931 (9.6%) among adults 40 years of age or older, with some modest variations between carriage surveys (Table 1). There was a direct strong, positive association between the prevalence of serotypes among carriers younger than 7 years of age and the prevalence of serotypes carried among adults 18–39 years of age (Spearman's  $r = 0.85$ , 95% confidence interval (CI): 0.80, 0.89) and adults 40 years of age or older (Spearman's  $r = 0.63$ , 95% CI: 0.53, 0.71) (Figure 1, Web Figure 1).

### Invasiveness patterns in children and adults

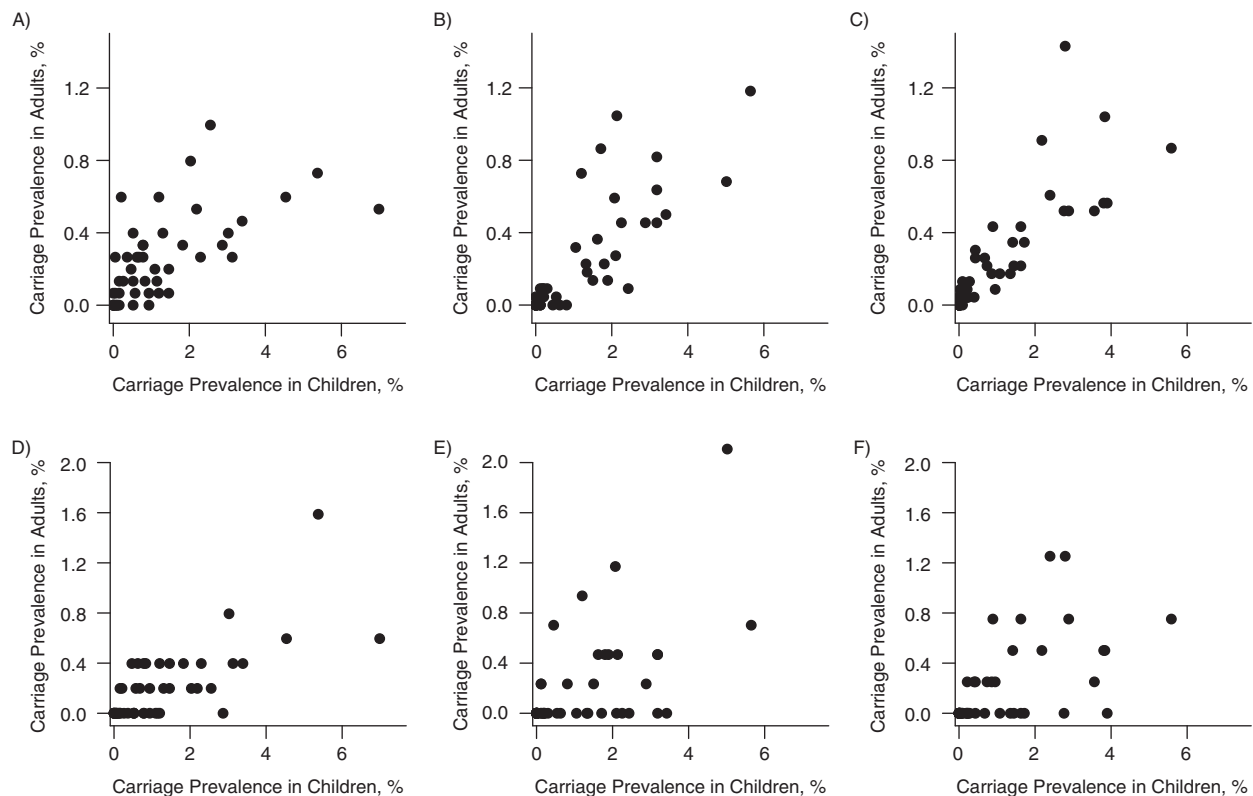
Despite differences in carriage prevalence between age groups, the invasiveness (disease incidence divided by carriage prevalence) of individual serotypes had similar rank orders across age groups (Figure 2, Web Figure 2). There was a moderate association between invasiveness in children and invasiveness in adults 18–39 years of age (Spearman's  $r = 0.50$ , 95% CI: 0.27, 0.68) and in adults 40 years old age or older (Spearman's  $r = 0.58$ , 95% CI: 0.36, 0.75). Serotypes 1, 7F, and 12F were among the most invasive serotypes

across all age groups. However, serotypes 1, 9V, and 18C were significantly more invasive in children than in adults, whereas serotypes 3, 6A/C, 8, and 11A were significantly more invasive in adults than in children (95% credible intervals for the differences between the estimates from adults and children did not include 0).

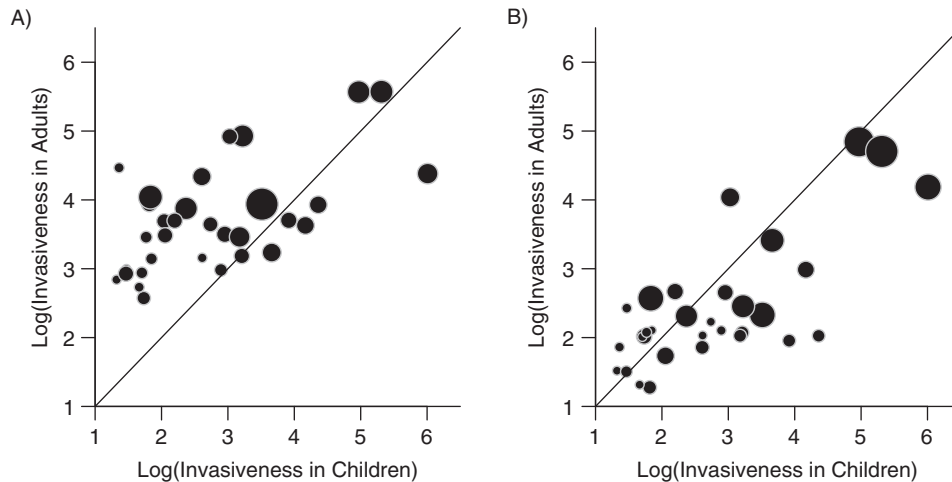
### Use of pediatric carriage and invasiveness data as a predictor of IPD in children and adults

Carriage and invasiveness data in adults are sparse, and we were interested in how carriage in children relates to disease in adults. Therefore, we considered whether carriage in children and invasiveness in children could be used to predict IPD patterns in adults. Models were fit to adult IPD data from 3 of the 4 available time periods and then used to predict the fourth time period based on the observed carriage data. When aggregating across all serotypes, the models captured the predicted number of cases of IPD for most time periods and age groups (Figure 3).

For the majority of serotypes and periods, the 95% credible interval for the predicted values included the observed values, but the credible intervals for individual serotypes were wide in some instances (Web Figures 3–5). There were strong associations between the serotype-specific observed and predicted numbers of cases in children younger than 7 years of age (Spearman's correlations of 0.84, 0.80, 0.71, and 0.54

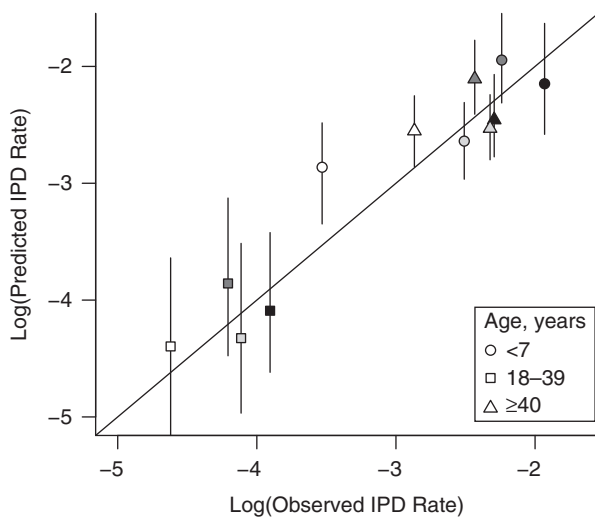


**Figure 1.** Comparison of carriage frequency among children younger than 7 years of age with that in adults 18–39 years of age (A–C) or that in adults 40 years of age or older (D–F), Navajo and White Mountain Apache populations, United States, 2000–2012. Each panel represents a different time period: early post-pneumococcal conjugate vaccine (PCV7) period (A, D), late post-PCV7 period (B, E), and early post-PCV13 period (C, F). Carriage data from adults during the pre-PCV7 period were not available. For serotype labels, see Web Figure 1.



**Figure 2.** Comparison of invasiveness patterns (incidence of disease divided by frequency of carriage) in children younger than 7 years of age and those in A) adults 18–39 years of age or B) adults 40 years of age or older, Navajo and White Mountain Apache populations, United States, 1995–2012. The size of the bubble is proportional to the precision of the estimates (i.e., larger bubbles indicate more confidence). The diagonal line denotes  $X = Y$ . For serotype labels, see Web Figure 2.

for each of the 4 time periods), adults 18–39 years of age (Spearman's correlations of 0.54, 0.46, 0.59, and 0.67), and adults 40 years of age or older (Spearman's correlations of 0.83, 0.76, 0.80, and 0.73).



**Figure 3.** The observed rate of invasive pneumococcal disease (IPD; cases per 1,000 persons in the population) in each age group and time period compared with the number of cases predicted using a regression model, Navajo and White Mountain Apache populations, United States, 1995–2012. The regression included serotype-specific carriage frequency in children, invasiveness in children, and a random effect for serotype. The estimates are based on a cross-validation approach in which the time period shown was held out during model fitting, and predictions are based on observed carriage prevalence during the time period. The colors indicate the time period (black, pre-pneumococcal conjugate vaccine (PCV); dark gray, early post-PCV7; light gray, late post-PCV7; and white, early post-PCV13).

Carriage prevalence and invasiveness were log-transformed in the Poisson regression model so that a regression coefficient of 1 would indicate that a 1-log increase in carriage prevalence or invasiveness was associated with a 1-log increase in IPD incidence, a value of between 0 and 1 would indicate that an increase in carriage in children was associated with a smaller increase in disease in adults, and a coefficient of 0 would indicate no association. For models fit to the post-PCV7 data, the coefficients for invasiveness were not significantly different from 1 in any of the age groups (among children younger than 7 years of age, Spearman's  $r = 1.10$ , 95% credible interval (CrI): 0.90, 1.34; among adults 18–39 years of age, Spearman's  $r = 0.96$ , 95% CI: 0.61, 1.45; and among adults 40 years of age or older, Spearman's  $r = 0.77$ , 95% CrI: 0.50, 1.01). The coefficients for carriage were not different from 1 for children (Spearman's  $r = 0.98$ , 95% CrI: 0.76, 1.24). In contrast, the coefficients were smaller among adults 18–39 years of age (Spearman's  $r = 0.63$ , 95% CrI: 0.35, 0.87) and adults 40 years of age or older (Spearman's  $r = 0.55$ , 95% CrI: 0.36, 0.74).

Among adults 40 years of age or older, there was also significant variation in disease rates between serotypes that was not explained by carriage in children or invasiveness in children. This is captured by the serotype-specific random effect in the model. In particular, serotypes 3, 8, and 12F consistently had random intercepts with 95% credible intervals that were above 0, indicating that they tended to cause more disease than could be explained based on the pediatric carriage prevalence and pediatric invasiveness (Web Figure 6). Likewise, serotypes 14 and 23F had negative intercepts, indicating that they caused less disease than expected (Web Figure 6).

#### Prediction of disease patterns in adults with or without risk factors

Finally, we considered whether the associations of IPD with carriage prevalence in children and with invasiveness

in children differed when predicting IPD among adults 40 years of age or older with or without comorbid conditions included in the Charlson comorbidity index. The fits of the models to the data were similar between groups (i.e., similar numbers of serotypes for which the 95% credible intervals contained the observed values) but were notably less accurate than for the unstratified analyses (Web Figures 7–9). As with the unstratified analysis, the random intercepts indicated that there was significant variation in disease rates between serotypes that was not explained by carriage in children or invasiveness in children. This unexplained variation among the serotypes was modestly, but not significantly, greater among those with comorbid conditions (Web Figure 10).

## DISCUSSION

With PCVs now being administered to children around the world, it is critical to understand how bacterial carriage patterns among children relate to disease patterns in adults. Much of the population-wide impact of using PCVs in children (i.e., total cases averted) results from indirect protection of adults (25). A key measure for understanding these impacts is the invasiveness ratio, which links carriage and disease. We found that despite differences in absolute pneumococcal carriage prevalence and disease rates between age groups, the invasiveness of the serotypes was similar across ages, and the rank orders of carriage prevalence by serotype among healthy children and adults were similar. Critically, carriage and invasiveness data from children can be used to estimate IPD patterns in adults. Combined with models previously developed by Nurhonen and Auranen (7), these results, along with transmission models, could help to forecast the impacts on disease rates in different age and risk groups of introducing new pneumococcal vaccine formulations (e.g., PCV15) in children.

These results build on those from previous studies in which investigators have examined the relationships of carriage frequency in children and invasiveness in children with disease rates in children (2, 7–9). The present study is unique in that it was explicitly focused on the link of carriage and invasiveness in children with disease patterns in adults. The regression models also included a random effect for serotype, which captured additional variation in IPD rates that was not explained by prevalence or invasiveness in children. Among adults 40 years of age or older, a number of serotypes had significant random effects, which indicates that these serotypes caused more or less IPD in older adults than would be expected based only on pediatric carriage and invasiveness (Web Figure 4). This might indicate that either the serotypes are transmitted more or less commonly among adults than among children or that adults are more or less susceptible to diseases caused by these serotypes than are children. Notably, the invasiveness of serotypes 3 and 8 was significantly higher in adults, and these serotypes had significant serotype-specific random intercepts. This suggests that the random intercepts could simply be capturing variations in invasiveness between age groups. Among children younger than 7 years of age and adults 18–39 years of age, there were no significant serotype-specific random effects (Web Figure 4). The random effects were notably small for the children younger

than 7 years of age, and this provides a check on the model structure—by definition, pediatric carriage multiplied by pediatric invasiveness should give pediatric IPD incidence, so we would not expect substantial additional variation associated with serotype.

Our analyses relied on 2 major assumptions: 1) that the invasiveness of a serotype is consistent over time and 2) that carriage in adults is directly related to carriage in children. The first assumption is supported by results from meta-analyses (26), and to date there has not been evidence of changes in invasiveness after the introduction of PCVs. If a new, highly invasive clone were to emerge, this would invalidate the relationship between carriage and disease for that serotype. Recent work demonstrates that noncapsular factors can influence the invasiveness of a strain (14). However, worldwide similarity in invasiveness patterns despite the global diversity of pneumococcal noncapsular factors suggests that the capsule itself (or some factor linked to the capsule) is the dominant determinant of invasiveness patterns at the population level. The assumption about the link between carriage in children and adults is supported by the correlation in serotype frequencies among pediatric and adult carriers (Figure 1). However, it is possible that in some institutional settings (e.g., nursing homes), transmission among adults is more important than transmission from children to adults. Likewise, some serotypes, such as serotype 1, could rely on transmission among sick adults or could be exclusively carried among adults. The serotype-specific random effects in our model should capture any such patterns.

The predictions were inaccurate for a few of the serotypes during some time periods. Notably, the serotypes for which the predictions were inaccurate tended to be the serotypes that exhibit strong multiyear epidemic cycles (e.g., 1 and 12F) (27). These dynamics challenge efforts to make predictions about vaccine impacts. Because the carriage data are often collected over shorter time periods than are the disease data (2), there can be mismatches between these data sets, which in turn result in inaccuracies in predictions based on the carriage data. Either adjusting for the dynamics in the disease data using regression models (28) (where possible) or conducting carriage sampling continuously over time could allow for more accurate matches between the carriage and disease data.

These studies were conducted among members of the Navajo Nation and White Mountain Apache tribe in the southwestern United States. This population has high rates of both nasopharyngeal carriage and invasive pneumococcal disease compared with the general population in the United States. Additionally, the age distribution of the Navajo and White Mountain Apache differs from that of the rest of the United States population, with shorter life expectancy and high pneumococcal disease rates in middle-aged adults. As a result, these findings might not be generalizable to other populations in the United States and Europe, where the risk factors for disease and age distributions differ. However, these findings might be applicable to other resource-poor settings that are now introducing the vaccine, including some countries in Africa or Asia.

We chose to report 95% credible intervals for the predictions from the models. Although this is a conventional cutoff, a 90% or 99% credible interval would provide narrower or

wider credible intervals and might influence whether or not the observed values fall within the credible intervals.

Our results are subject to several additional limitations. We assumed that the IPD case patients with missing serotype data had a same serotype distribution that was the same as the observed serotype distribution during the same time period. This information is missing because of sample processing issues rather than any systematic bias, but if this assumption did not hold, it could bias the study results. These analyses were not adjusted for possible effects of PPV23, a vaccine with high uptake in this population; PPV23 use would effectively reduce the invasiveness of the 23 serotypes included in that vaccine if it were effective against all of the included serotypes. Such an effect would likely be captured by the serotype-specific random effects, given that uptake was consistent over time.

In conclusion, we have demonstrated that serotype-specific carriage and invasiveness patterns are broadly similar between children and adults and that carriage prevalence in children, invasiveness in children, and a serotype-specific random effect can be used to predict disease incidence in adults for many serotypes in the post-PCV era. These findings will help to understand and predict how the use of pneumococcal conjugate vaccines in children are likely to impact disease rates in adults.

Additional Web Material: R Analysis code and data:

1. R code and data for running the cross-validation analysis in JAGS. Random integers were added to the carriage and invasive pneumococcal disease data for privacy reasons
2. R code, sample data (with random integers added), and instructions for generating predictions based on pediatric carriage data

## ACKNOWLEDGMENTS

Author affiliations: Department of Epidemiology of Microbial Diseases, School of Public Health, Yale University, New Haven, Connecticut (Daniel M. Weinberger); Center for American Indian Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland (Lindsay R. Grant, Robert C. Weatherholtz, Katherine L. O'Brien, Laura L. Hammitt); Department of Biostatistics, School of Public Health, Yale University, New Haven, Connecticut (Joshua L. Warren); and International Vaccine Access Center, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland (Katherine L. O'Brien, Laura L. Hammitt).

D.M.W. was supported by grant R56AI110449 from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, grant P30AG021342 from the National Institute on Aging, National Institutes of Health (scholar at the Claude D. Pepper Older Americans Independence Center at Yale University School of Medicine), grant UL1 TR000142 from the National Institutes of Health (Clinical and Translational Science Award to Yale University), a grant from the Bill and Melinda Gates Foundation, and an investigator-initiated research grant from Pfizer. The collection of the original data used in this analysis was supported by the Native American Research Centers for Health (grant

U26IHS300013/03), the Centers for Disease Control and Prevention, the Trasher Research Fund, the National Institute on Minority Health and Health Disparities, National Institutes of Health (grant R01MD004011), and Pfizer (protocol number: 6096A1-4013). L.R.G. and R.W. received research support in the past 3 years from the National Institutes of Health. K.L.O. received pneumococcal research grant funding in the past 3 years from the National Institutes of Health, The Bill and Melinda Gates Foundation, and Gavi, the Vaccine Alliance.

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the Indian Health Service.

Conflict of interest: D.M.W., L.L.H., and K.L.O. have participated in a Scientific Input Engagement for Merck. D.M.W. has received consulting fees from Pfizer and Affinivax. L.R.G. and R.W. have received research support in the past 3 years from GlaxoSmithKline and Pfizer. K.L.O. received pneumococcal research grant funding in the past 3 years from GlaxoSmithKline. L.L.H. received research funding from GlaxoSmithKline and Pfizer.

## REFERENCES

1. Tomczyk S, Bennett NM, Stoecker C, et al. Use of 13-valent pneumococcal conjugate vaccine and 23-valent pneumococcal polysaccharide vaccine among adults aged  $\geq 65$  years: recommendations of the Advisory Committee on Immunization Practices (ACIP). *Morb Mortal Wkly Rep*. 2014;63(37):822–825.
2. Weinberger DM, Bruden DT, Grant LR, et al. Using pneumococcal carriage data to monitor postvaccination changes in invasive disease. *Am J Epidemiol*. 2013;178(9):1488–1495.
3. Feikin DR, Kagucia EW, Loo JD, et al. Serotype-specific changes in invasive pneumococcal disease after pneumococcal conjugate vaccine introduction: a pooled analysis of multiple surveillance sites. *PLoS Med*. 2013;10(9):e1001517.
4. Walter ND, Taylor TH Jr, Dowell SF, et al. Holiday spikes in pneumococcal disease among older adults. *N Engl J Med*. 2009;361(26):2584–2585.
5. Bogaert D, de Groot R, Hermans PWM. *Streptococcus pneumoniae* colonisation: the key to pneumococcal disease. *Lancet Infect Dis*. 2004;4(3):144–154.
6. Simell B, Auranen K, Käyhty H, et al. The fundamental link between pneumococcal carriage and disease. *Expert Rev Vaccines*. 2012;11(7):841–855.
7. Nurhonen M, Auranen K. Optimal serotype compositions for pneumococcal conjugate vaccination under serotype replacement. *PLoS Comput Biol*. 2014;10(2):e1003477.
8. Weinberger DM, Harboe ZB, Flasche S, et al. Prediction of serotypes causing invasive pneumococcal disease in unvaccinated and vaccinated populations. *Epidemiology*. 2011;22(2):199–207.
9. Shea KM, Weycker D, Stevenson AE, et al. Modeling the decline in pneumococcal acute otitis media following the introduction of pneumococcal conjugate vaccines in the US. *Vaccine*. 2011;29(45):8042–8048.
10. Yildirim I, Hanage WP, Lipsitch M, et al. Serotype specific invasive capacity and persistent reduction in invasive pneumococcal disease. *Vaccine*. 2010;29(2):283–288.

11. Brueggemann AB, Spratt BG. Geographic distribution and clonal diversity of *Streptococcus pneumoniae* serotype 1 isolates. *J Clin Microbiol*. 2003;41(11):4966–4970.
12. Hyams C, Yuste J, Bax K, et al. *Streptococcus pneumoniae* resistance to complement-mediated immunity is dependent on the capsular serotype. *Infect Immun*. 2010;78(2):716–725.
13. Hyams C, Trzcinski K, Camberlein E, et al. *Streptococcus pneumoniae* capsular serotype invasiveness correlates with the degree of factor H binding and opsonization with C3b/iC3b. *Infect Immun*. 2013;81(1):354–363.
14. Browall S, Norman M, Tångrot J, et al. Intraclonal variations among *Streptococcus pneumoniae* isolates influence the likelihood of invasive disease in children. *J Infect Dis*. 2014; 209(3):377–388.
15. Trotter CL, Waight P, Andrews NJ, et al. Epidemiology of invasive pneumococcal disease in the pre-conjugate vaccine era: England and Wales, 1996–2006. *J Infect*. 2010;60(3): 200–208.
16. Watt JP, O'Brien KL, Benin AL, et al. Risk factors for invasive pneumococcal disease among Navajo adults. *Am J Epidemiol*. 2007;166(9):1080–1087.
17. Weatherholtz R, Millar EV, Moulton LH, et al. Invasive pneumococcal disease a decade after pneumococcal conjugate vaccine use in an American Indian population at high risk for disease. *Clin Infect Dis*. 2010;50(9):1238–1246.
18. Scott JR, Millar EV, Lipsitch M, et al. Impact of more than a decade of pneumococcal conjugate vaccine use on carriage and invasive potential in Native American communities. *J Infect Dis*. 2012;205(2):280–288.
19. Millar EV, Watt JP, Bronsdon MA, et al. Indirect effect of 7-valent pneumococcal conjugate vaccine on pneumococcal colonization among unvaccinated household members. *Clin Infect Dis*. 2008;47(8):989–996.
20. O'Brien KL, Millar EV, Zell ER, et al. Effect of pneumococcal conjugate vaccine on nasopharyngeal colonization among immunized and unimmunized children in a community-randomized trial. *J Infect Dis*. 2007;196(8):1211–1220.
21. Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis*. 1987;40(5): 373–383.
22. Harrison XA. Using observation-level random effects to model overdispersion in count data in ecology and evolution. *PeerJ*. 2014;2:e616.
23. Plummer M. rjags: Bayesian graphical models using MCMC. R package version 3-14. 2014.
24. R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2013.
25. Moore MR, Link-Gelles R, Schaffner W, et al. Effect of use of 13-valent pneumococcal conjugate vaccine in children on invasive pneumococcal disease in children and adults in the USA: analysis of multisite, population-based surveillance. *Lancet Infect Dis*. 2015;15(3):301–309.
26. Brueggemann AB, Peto TE, Crook DW, et al. Temporal and geographic stability of the serogroup-specific invasive disease potential of *Streptococcus pneumoniae* in children. *J Infect Dis*. 2004;190(7):1203–1211.
27. Harboe ZB, Benfield TL, Valentiner-Branth P, et al. Temporal trends in invasive pneumococcal disease and pneumococcal serotypes over 7 decades. *Clin Infect Dis*. 2010; 50(3):329–337.
28. Harboe ZB, Dalby T, Weinberger DM, et al. Impact of 13-Valent pneumococcal conjugate vaccination in invasive pneumococcal disease incidence and mortality. *Clin Infect Dis*. 2014;59(8):1066–1073.