Temporal and Geographic Stability of the Serogroup-Specific Invasive Disease Potential of \textit{Streptococcus pneumoniae} in Children

\author{Angela B. Brueggemann,\textsuperscript{1,2} Timothy E. A. Peto,\textsuperscript{1} Derrick W. Crook,\textsuperscript{1} Jay C. Butler,\textsuperscript{5} Karl G. Kristinsson,\textsuperscript{4} and Brian G. Spratt\textsuperscript{3}}

\textsuperscript{1}Academic Department of Microbiology and Infectious Disease and \textsuperscript{2}Department of Public Health and Primary Care, University of Oxford, Oxford, and \textsuperscript{3}Department of Infectious Disease Epidemiology, Imperial College London, St. Mary’s Hospital, London, United Kingdom; \textsuperscript{4}Clinical Microbiology, Landspitali University Hospital, Reykjavik, Iceland; \textsuperscript{5}Arctic Investigations Program, Centers for Disease Control and Prevention, Anchorage, Alaska

A meta-analysis study design was used to analyze 7 data sets of invasive and carriage pneumococcal isolates recovered from children, to determine whether invasive disease potential differs for each serotype and, if so, whether it has changed over time or differs geographically. Serotype- and serogroup-specific odds ratios (ORs) were calculated for each study and as a pooled estimate, with use of serotype 14 as the reference group. ORs varied widely: the serotypes with the highest ORs (1, 5, and 7) were 60-fold more invasive than those with the lowest ORs (3, 6A, and 15). There was a significant inverse correlation between invasive disease and carriage prevalence for the serotypes that we considered, which implies that the most invasive serotypes and serogroups were the least commonly carried and that the most frequently carried were the least likely to cause invasive disease. There was no evidence of any temporal change or major geographical differences in serotype- or serogroup-specific invasive disease potential.

\textit{Streptococcus pneumoniae} is one of the leading bacterial causes of infection worldwide, and it is responsible for illnesses ranging from mild upper respiratory infections to life-threatening invasive diseases, such as bacteremia and meningitis. The incidence of invasive pneumococcal disease in young children residing in the United States and Europe is 8–75 cases/100,000 population/year, whereas the incidence in young children residing in developing countries is several times higher, 100 to >500 cases/100,000 population/year. Populations such as American Navajo Indians, Australian Aboriginals, and Alaska Natives also have high rates that may be >1000 cases/100,000 population/year [1, 2]. Major efforts have been undertaken in recent years to develop and implement a vaccine for the prevention of pneumococcal disease in young children, the age group most at risk of disease. The heptavalent conjugate pneumococcal vaccine was licensed for use in the United States in 2000 and is currently proving to be remarkably effective at preventing invasive pneumococcal disease in vaccinated children. Furthermore, there is evidence of significant herd immunity as a result of widespread vaccination, which has resulted in a concomitant reduction in invasive disease rates in unvaccinated older children and adults [3].

The major ecological niche of the pneumococcus is the nasopharynx of healthy children. Numerous articles have reported pneumococcal carriage rates in healthy children ranging from <10% during the weeks after birth to 100% during the first few years of life, and nearly all children are believed to have at least 1 episode of carriage during early childhood [4]. However, although most healthy children have carried a pneumococcus, very few develop invasive disease; therefore, invasive disease is a relatively rare event. The vast majority of the pneumococcal population resides harmlessly in
the nasopharynx of healthy children, and this carriage state is important for both the pneumococcus and the host. For example, natural selection applied by the use of antibiotics or selection for strains with increased transmission to new hosts occurs among the carriage population, which occasionally results in strains that have acquired increased resistance to antibiotics or increased ecological success; also of importance is that the invasion of the bloodstream is preceded by the acquisition and colonization of the host mucosal surface by the pneumococcus. The morbidity and mortality associated with invasive pneumococcal disease is significant, and the prevention of disease is of considerable public health importance. Conjugate vaccines prevent invasive disease caused by vaccine serotypes but also reduce the carriage of vaccine serotypes and increase the carriage of nonvaccine serotypes (serotype replacement) [5–8]. We need to understand vaccine-induced changes in the ecology of S. pneumoniae and the relationship between the organism’s niche in the nasopharynx and invasive disease, to gain a better understanding of which pneumococci are most capable of causing disease in a population with high rates of immunization with the pneumococcal conjugate vaccine.

S. pneumoniae has a number of virulence factors, perhaps the most important of which is a polysaccharide capsule. Chemical and antigenic differences in this capsule result in 90 different capsular types or serotypes, ~15 of which cause the majority of invasive disease [9, 10]. The capsule is the target for the protective immunity induced by conjugate vaccines, and the selection of serotypes to be included in these vaccines was guided by the rank order of serotypes causing invasive disease, under the assumption that the most prevalent serotypes must be the most invasive. However, a serotype may rank high in the list of invasive serotypes simply because children are very frequently exposed to that serotype and occasionally develop disease, not because that serotype is particularly invasive. Conversely, a highly invasive serotype may be much more likely to cause disease after acquisition, but, if it is infrequent in the population, then it will appear to be less invasive simply because it is lower in the rank order of serotypes causing disease.

Only 2 studies have attempted to relate the prevalence of invasive pneumococcal disease caused by different serotypes in a community to the extent of exposure of that population to these serotypes [11, 12]. One of these studies characterized, at both a serotypic and a genotypic level, of a collection of invasive and carriage pneumococci recovered from a population of young children in Oxford, United Kingdom [12]. The major carriage and invasive clones of pneumococci were identified, and odds ratios (ORs) were calculated to estimate the invasive disease potential of each serotype and major clone. The data suggested that individual serotypes varied in invasive disease potential—for example, serotypes such as 1, 14, and 18C were significantly more likely to cause invasive disease than to be carried in the nasopharynx, whereas serotype 23F caused significantly less invasive disease than had been expected from its prevalence in carriage. A strong association between genotype and capsular type and a lack of heterogeneity in the invasiveness of different clones of the same serotype suggested that capsular type may be a more important determinant of the likelihood of invasive disease than is the overall genotype.

A major strength of the Oxford study was that the isolates causing invasive disease could be related to those that were circulating among Oxfordshire children of similar age during a defined period of time. A limitation of the study was that it represented data from a single population of children, and the results may not be generalizable to other populations. Furthermore, cross-sectional carriage data were used as a proxy for acquisition data, and differences in the duration of carriage could have contributed to the differences in the estimates of invasive disease potential [12]. Therefore, the aim of the present study was to use as many data sets as possible to explore the observation that pneumococcal serotypes vary in their potential to cause invasive disease (i.e., the potential of a pneumococcus expressing a particular capsular type to invade normally sterile body fluids such as blood and cerebrospinal fluid [CSF]) and to explore whether invasive disease potential differs temporally or geographically. To this end, we studied large collections of similarly well-sampled carriage and invasive pneumococci from young children residing in Alaska and Iceland, and a meta-analysis was performed by use of data sets from Oxford, Alaska, and Iceland and data from studies previously published in the medical literature.

**METHODS AND MATERIALS**

**Literature search and inclusion criteria.** For the purposes of the meta-analysis, we were only interested in studies that characterized invasive and carriage pneumococcal isolates from young children residing in the same geographic region during approximately the same period of time. The serotype or serogroup distribution for both invasive and carriage pneumococcal isolates had to be given in a table or figure such that it was possible to extract the raw numbers of isolates in each group. A MEDLINE search of the literature written in English was performed by use of combinations of the terms “streptococcus pneumoniae” or “pneumococcus” and “invasive disease” or “carriage,” and then the search was expanded by use of the bibliographic reference lists from the relevant studies identified via the MEDLINE search. Numerous studies were identified that reported only the serotype distribution of pneumococci causing invasive disease among children or only the serotype distribution of pneumococcal isolates recovered from healthy children; only 10 studies reported the serotype distribution of both invasive and carriage pneumococcal isolates. From these 10 studies, 5 were excluded for the following rea-
The aim of the study was to calculate and compare serotype- and serogroup-specific ORs for invasive disease for each individual study and for the pooled data. There were specific serotypes and serogroups of interest, including the serotypes included in the current 7-valent conjugate vaccine (4, 6B, 9V, 14, 18C, 19F, and 23F), the serotypes that are included in higher-valency conjugate vaccines (1, 3, 5, and 7F), and the nonvaccine serotypes or serogroups that were present in sufficient quantities for analysis. Only 4 studies contained serotype data (Toronto, Oxford, Alaska, and Iceland), and 2 of these contained serotype data for only a proportion of their isolate collection (Toronto and Iceland). Therefore, for purposes of the meta-analysis, serogroup data were used, and, when a serogroup included multiple serotypes, the serotype data were pooled into the serogroup (e.g., serotypes 19A and 19F were pooled into serogroup 19 for the analysis). There were sufficient numbers of pneumococcal isolates for an analysis of serogroups 1, 3–9, 14, 15, 18, 19, 23, 33, and 38.

Calculation of the OR. We calculated a serogroup-specific OR to estimate the probability of invasive disease due to each individual serogroup, using the standard OR equation (see below). The selection of the reference group to which each serogroup was compared was considered carefully, to ensure that the resulting estimate was a robust measure of invasiveness. One option was to compare serogroup X with all non-X serogroups within each study, as has been done elsewhere [11, 12]. However, because the quantity of each serogroup varied greatly between studies, and because some prevalent serogroups contain multiple serotypes (e.g., 19A and 19F) that, in most studies, were not stratified by serotype, we elected to use a fixed reference set instead. Serotype 14 was chosen as the referent serotype because there was no evidence of heterogeneity in the combined serotype 14 data, it exists as a single serotype (i.e., there are no subtypes), it is among the most prevalent invasive and carriage serotypes in all of these studies, and it has been shown in a previous study to be among the more invasive pneumococcal serotypes [12]. Thus, the serogroup-specific ORs for each study were calculated by use of the proportion of invasive and carriage serotype 14 pneumococci within that study as the referent group and by use of the following equation: OR = (ad)/(bc), where a is the number of invasive serotype X isolates, b is the number of carriage serotype X isolates, c is the number of invasive serotype 14 isolates, and d is the number of carriage serotype 14 isolates. Finally, a pooled estimate of the serogroup-specific OR was calculated by combining the individual ORs by use of conventional Mantel-Haenszel statistics. Therefore, each OR may be defined as an estimate of serotype- or serogroup-specific invasive disease potential, relative to the invasive disease potential of serotype 14.

Statistical analyses. The ORs and Mantel-Haenszel 95% confidence intervals (CIs) were calculated by use of the Stata software package (version 8; Stata). ORs for the serogroups listed above were calculated for each individual study and for the pooled estimate by use of the “metan” command. The validity of the meta-analysis was dependent on the lack of statistically significant heterogeneity between the serotype- and serogroup-specific ORs for individual data sets, which, if present, would have biased the pooled OR estimates when multiple data sets were combined. Therefore, the data for each serotype or serogroup were analyzed for statistical evidence of heterogeneity before the calculation of the OR estimates; because of multiple comparisons, a significance level of .01 was used as evidence of heterogeneity.

RESULTS

Table 1 describes the characteristics of the 7 data sets included in the meta-analysis. The studies were nearly evenly distributed during 1975–2002 and resulted in an analysis of 769 invasive and 4927 carriage pneumococcal isolates. All of the invasive isolates were recovered from blood, CSF, or other normally sterile body fluids. Carriage isolates were obtained via nasal or nasopharyngeal swabs, except in the Alabama study, where isolates were recovered from throat swabs. All of the children from whom the samples were obtained were <18 years old, although the majority were <6 years old.

We tested for between-study heterogeneity in the calculated study-specific ORs, and 2 serogroups, 6 and 7, revealed evidence of significant heterogeneity (P = .005 and P = .009, respectively). Because 4 of the studies did include data for the individual serotypes within serogroup 6 (6A and 6B; table 1), we calculated a serotype-specific OR for each of those 4 studies, and this analysis resolved the heterogeneity. That is, there was no evidence of significant heterogeneity within either the serotype 6A or 6B data when they were analyzed as serotypes rather than as part of pooled serogroup 6; therefore, a pooled serotype 6A OR estimate
Table 1. Characteristics of each study included in a meta-analysis of the serogroup-specific invasive disease potential of *Streptococcus pneumoniae*.

<table>
<thead>
<tr>
<th>Study data</th>
<th>Study locationa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at time of sampling: invasive/carriage, years</td>
<td>&lt;6/GS</td>
</tr>
<tr>
<td>Serotype and serogroupb</td>
<td>1 4/2</td>
</tr>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>6c</td>
</tr>
<tr>
<td></td>
<td>6A</td>
</tr>
<tr>
<td></td>
<td>6B</td>
</tr>
<tr>
<td></td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>38</td>
</tr>
<tr>
<td>Total collection</td>
<td>105/235</td>
</tr>
</tbody>
</table>

NOTE. GS, grammar school; NA, not available; PNG, Papua New Guinea; ... serotype/serogroup absent from the collection.

a References for each study: Alabama [13], PNG [11], Kenya [14], Toronto [15], Oxford [12], Iceland (K.G.K. unpublished data), and Alaska (J.C.B. unpublished data).
b No. of isolates, invasive/carriage.
c The no. of serogroup 6 isolates in each study is indicated, and the serotype (6A or 6B) was determined in 4 studies: Toronto, Iceland (8 isolates were not serotyped), Oxford, and Alaska.

and a pooled serotype 6B OR estimate were calculated. Six of the studies reported serogroup 7 data, but the sample size of this serogroup was small in most studies (Oxford and Alaska, n = 3; Alabama and Kenya, n = 5; Iceland, n = 22; and Papua New Guinea, n = 32). Furthermore, serotype data for serogroup 7 were only available from Oxford (n = 3, all 7F) and Alaska (n = 3, all 7C); therefore, a serotype-specific analysis such as that performed for serotypes 6A and 6B was not possible. Thus, the study-specific ORs and the pooled OR estimate for serogroup 7 should be interpreted with caution.

Figures 1 and 2 display the study-specific OR estimates for the 7 serogroups that contained the serotypes included in the heptavalent conjugate pneumococcal vaccine (figure 1) and the OR estimates for other serotypes and serogroups not included in the current conjugate vaccine (figure 2). Figure 3 displays the calculated pooled estimate ORs in the meta-analysis, plotted as the OR of each serogroup versus the prevalence of that serogroup in carriage. The reference group for the pooled estimate OR was all invasive (n = 201) and carriage (n = 273) serotype 14 isolates. The pooled estimate ORs indicated that serotypes 1 and 7 were 2–6 times more invasive than serotype 14; serotype 4 was equivalent to serotype 14 in invasiveness; and all other serogroups were less invasive than serotype 14, with ORs of 0.1–0.6. There was also an inverse relationship between invasiveness and carriage prevalence—that is, increased OR estimates correlated with a decreased prevalence in carriage (P < .03).

**DISCUSSION**

The main purpose of the present study was to measure the differences in the ability of pneumococcal serogroups to cause invasive disease and to determine whether this invasive disease potential differs temporally or geographically. We used a meta-analysis study design to compare pneumococcal data sets from 7 different locations and measured the invasive disease potential by calculating serogroup-specific ORs for each study and the pooled data. The studies included in the meta-analysis spanned a 27-year period and were from several different types of geographic locations and populations: low-income countries (Kenya and Papua New Guinea); medium to large cities (Birmingham, Alabama; Toronto, Ontario; and Oxford, England), and unique populations (Iceland, a geographically isolated island with a population of <300,000 people; and Alaska, a noncontiguous US state with a population of <700,000, of which 20% are Alaska Native or American Indian, a population that has previously been shown to be at increased risk of invasive pneumococcal disease [16]).
Figure 1. Estimates of serotype- and serogroup-specific invasive disease potential for each study in the meta-analysis, relative to the invasive disease potential of serotype 14. Odds ratios (ORs) and 95% confidence intervals for heptavalent vaccine or vaccine-related serogroups and serotypes, including \( P \) values for the test of between-study heterogeneity, are shown. \( P \geq .01 \) indicates that a pooled estimate of the serogroup-specific OR is statistically valid, because there is no significant heterogeneity in the pooled data sets for the serogroup. ALA, Alabama; ASK, Alaska; ICE, Iceland; KEN, Kenya; OXF, Oxford; PNG, Papua New Guinea.
Figure 2. Study-specific odds ratios (ORs) and 95% confidence intervals (CIs) for serotypes and serogroups not included in the heptavalent pneumococcal conjugate vaccine but proposed for future vaccines (1, 3, 5, and 7F) or present in sufficient numbers for the present analysis (8, 15, 33, and 38). The OR for serotype 1 from the Papua New Guinea data set is 1742.57 (95% CI, 5.7–535,202.3); only the lower bound of the 95% CI is shown. ALA, Alabama; ASK, Alaska; KEN, Kenya; OXF, Oxford.
Figure 3. Correlation between pooled estimate odds ratios (ORs) and the prevalence of each serogroup among the pneumococcal carriage population. The relationship between the 2 measurements is linear and statistically significant ($P = .03$). The upper bounds of the confidence intervals for serotypes 1 and 5 are beyond the scale of the graph but are 10.7 and 21.9, respectively. Heptavalent conjugate vaccine serotypes and serogroups are indicated by black diamonds and vaccine-related or nonvaccine serotypes and serogroups by white diamonds.
The results of the meta-analysis supported the conclusion that pneumococcal serotypes and serogroups differ in invasiveness. The OR estimates varied widely (0.1–6.1), and those serotypes that had the highest ORs (1, 5, and 7) were ∼60-fold more invasive than those with the lowest ORs (serotypes 3 and 6A and serogroup 15). An OR >1 implies that the serogroup was more invasive than serotype 14. Because this reference serotype is also one of the most invasive serotypes, the high ORs of serotypes such as 1, 5, and 7 were particularly impressive. The upper limits of the CIs were very high for these serotypes, because they are very infrequently carried, even in locations where these serotypes are prevalent in invasive disease [17, 18].

An important result of the meta-analysis was that there was a statistically significant inverse relationship between invasive disease potential and carriage prevalence: the most invasive serogroups were the least commonly carried, whereas the serogroups most prevalent in carriage were the least likely to cause invasive disease. This could be partially related to differences in serotype- and serogroup-specific carriage duration. Preliminary data based on monthly swab samples from healthy children suggested that the serotypes that are the least invasive—such as 6B, 19F, and 23F—are also those that persist in the nasopharynx for long periods of time, whereas highly invasive serotypes—such as 7F and 14—are carried for shorter periods of time [4], and it is known that serotype 1 is almost never carried [5, 17]. However, the preliminary data indicate that there may only be a 3-fold difference in carriage duration, yet there is a 60-fold difference in invasive disease potential, which suggests that it is unlikely that the differences in carriage duration alone could account for the observed differences in invasive disease potential.

There are 2 other potential limitations of the meta-analysis to consider. One is that age may be associated with differences in serotype- and serogroup-specific invasive disease potential—that is, certain serotypes and serogroups may be more likely to cause meningitis in the youngest children, whereas others may be more commonly associated with bacteremia in older children. Although the age-stratified data necessary for such an analysis were not available for most of the data sets in the meta-analysis, we were unable to find any significant differences in the serotype distribution of the invasive isolates by age group (when stratified by age groups of <1, 1–2, and ≥2 years) in the data sets from Oxford, Iceland, and Alaska. Another potential limitation is that most of the studies in the meta-analysis did not provide serotype data; thus, it was necessary to analyze the data by serogroups, which could have masked serotype-specific differences in invasive disease potential (e.g., differences in invasiveness between serotypes 19A and 19F may not be apparent when an estimate of serogroup 19 invasive disease potential is calculated). However, the lack of statistically significant between-study heterogeneity among the pooled data sets suggests that neither age effects nor serotype-specific differences in invasiveness, if present, are significantly influencing the data and conclusions of the meta-analysis.

Historically, it was assumed that invasiveness correlated with rank order in invasive disease prevalence, but the results of the present meta-analysis support the view that estimating disease potential out of context with the level of exposure to each serogroup in a healthy population results in a biased interpretation of invasiveness. All of the studies included in the meta-analysis sampled invasive and carriage isolates from populations of children from the same community and during similar time periods, and the analysis revealed that serotypes often thought to be particularly invasive because they are high in the rank order of serotypes from invasive disease are actually much less invasive than other serotypes.

The lack of between-study heterogeneity among the data sets indicated that there was no evidence that serogroup-specific invasive disease potential differed by geographical location or has changed over the 27-year period during which the studies were undertaken. The lack of heterogeneity makes the meta-analysis statistically valid, but, more importantly, it suggests that serotype or serogroup captures a biological characteristic of the pneumococcus related to invasive disease potential. Whether the serotype or serogroup is the primary determinant in causing invasive disease remains to be determined; however, the present results support the hypothesis that serotype or serogroup is an important determinant of invasiveness or is a good marker for a genetic background that is particularly equipped to cause invasive disease. Capsular type is the only factor related to invasive disease potential that is under consideration in the current context, although there are many other factors that are likely to influence pneumococcal invasiveness and the severity of resulting disease. Determining which bacterial factors are most closely associated with the ability of a pneumococcus strain to invade and proliferate in normally sterile environments, such as blood and meninges, requires experimental study, but the present data should help guide the design of such studies and the selection of serotypes and serogroups to investigate.

Conjugate vaccines of higher valency will cover most of the invasive pneumococcal serotypes, and attention needs to be focused on whether there are nonvaccine serotypes that are not particularly invasive but that could cause considerable amounts of invasive disease within a vaccinated population. The nasopharynx is the major ecological niche for the pneumococcus, and, although the most commonly carried serotypes (e.g., 6B, 19F, and 23F) are not particularly invasive, because of their high prevalence in carriage they cause a large proportion of invasive disease cases worldwide and are appropriately included in the heptavalent and higher-valency vaccines. The question then arises: if serotype replacement occurs in vacci-
nated populations and results in high levels of carriage and transmission of some nonvaccine serotypes, will significant levels of disease caused by these nonvaccine serotypes also occur, even though they are not particularly invasive? At present, we know little about the long-term effects of pneumococcal conjugate vaccines on carriage and transmission or the extent to which vaccine serotypes are replaced by nonvaccine serotypes, and which of the latter may come to predominate among vaccinated children. High levels of carriage of nonvaccine serotypes should not be a problem if they are very poorly invasive, but there are serotypes, such as 3 (which has been proposed for 11-valent vaccines), 8, 33, and 38, that are reasonably prevalent among carriage serotypes and serogroups and show OR estimates of invasiveness that are similar to serotypes such as 6B, 19, and 23, which cause substantial levels of disease because of their high levels of carriage.

The first epidemiological report after the implementation of the heptavalent conjugate vaccine in the United States indicated that, although there was a significant decrease in invasive disease overall, there was also a 27% increase in invasive disease caused by nonvaccine serotypes \((n = 58\) cases, 15.7 estimated cases/100,000 population) [3]. This was not a statistically significant increase, but it suggests that changes in the pneumococcal population may already be taking place in the United States. Further support was provided in a recently published article from Pennsylvania [19] in which investigators reported a significant change in the distribution of pneumococcal serotypes recovered from patients with acute otitis media immediately after the introduction of the heptavalent conjugate vaccine. Changes in the carriage population may result in greater changes in the serotypes and serogroups causing acute otitis media than to those causing invasive disease, because there is evidence that there is less variation in the ability of strains and serotypes to cause acute otitis media than invasive disease [20]. Thus, changes among pneumococci causing acute otitis media, such as those described in the recent Pennsylvania study, may be an indirect measurement of vaccine-induced changes occurring in the pneumococcal carriage population. Such changes in carriage may ultimately have some effect on the serotypes causing invasive disease, particularly if a few nonvaccine serotypes with some ability to cause invasive disease become very commonly carried, and strongly emphasizes the need for population-based surveillance before and after vaccine implementation in those countries considering licensure of the conjugate pneumococcal vaccine.

References


